What About Short Run?

by

Lai Xu

Department of Economics
Duke University

Date: _____

Approved:

_____
Tim Bollerslev, Supervisor

_____
George Tauchen

_____
Andrew Patton

_____
Shakeeb Khan

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2014

Abstract

What About Short Run?

by

Lai Xu

Department of Economics
Duke University

Date: _____

Approved:

_____
Tim Bollerslev, Supervisor

_____
George Tauchen

_____
Andrew Patton

_____
Shakeeb Khan

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Economics in the Graduate School of Duke University
2014

Abstract

This dissertation explores issues regarding the short-lived temporal variation of the equity risk premium. In the past decade, the equity risk premium puzzle is resolved by many competing consumption-based asset pricing models. However, before Bollerslev et al. (2009), the return predictability as an outcome of such models has limited empirical support in the short-run. Nowadays, there has been a consensus of the literature that the short-run equity return's predictability is intimately linked with the variance risk premium—the difference between options-implied and actual realized variation measures.

In this work, I continue to argue the importance of the short-lived components in the equity risk premium. Specifically, I first provide simulation evidence of the strong return predictability based on the variance risk premium in the U.S. aggregate market, and document new empirical findings in the international setting. Then I attempt to use a structural macro-finance model to guide through the predictability estimation with much more efficiency gain. Finally I decompose the equity risk premium into two short-lived parts — tail risk and diffusive risk — and propose a semi-parametric estimation method for each part. The results are arranged in the following order.

Chapter 1 of the dissertation is co-authored with Tim Bollerslev, James Marrone and Hao Zhou. In this chapter, we demonstrate that statistical finite sample biases cannot "explain" this apparent predictability in U.S. market based on variance risk

premium. Further corroborating the existing evidence of the U.S., we show that country specific regressions for France, Germany, Japan, Switzerland, the Netherlands, Belgium and the U.K. result in quite similar patterns. Defining a "global" variance risk premium, we uncover even stronger predictability and almost identical cross-country patterns through the use of panel regressions.

Chapter 2 of the dissertation is co-authored with Tim Bollerslev and Hao Zhou. In this chapter, we examine the joint predictability of return and cash flow within a present value framework, by imposing the implications from a long-run risk model that allow for both time-varying volatility and volatility uncertainty. We provide new evidences that the expected return variation and the variance risk premium positively forecast both short-horizon returns *and* dividend growth rates. We also confirm that dividend yield positively forecasts long-horizon returns, but that it does not help in forecasting dividend growth rates. Our equilibrium-based "structural" factor GARCH model permits much more accurate inference than univariate regression procedures traditionally employed in the literature. The model also allows for the direct estimation of the underlying economic mechanisms, including a new volatility leverage effect, the persistence of the latent long-run growth component and the two latent volatility factors, as well as the contemporaneous impacts of the underlying "structural" shocks.

In Chapter 3 of the dissertation, I develop a new semi-parametric estimation method based on an extended ICAPM dynamic model incorporating jump tails. The model allows for time-varying, asymmetric jump size distributions and a self-exciting jump intensity process while avoiding commonly used but restrictive affine assumptions on the relationship between jump intensity and volatility. The estimated model implies that the average annual jump risk premium is 6.75%. The model-implied jump risk premium also has strong explanatory power for short-to-medium run aggregate market returns. Empirically, I present new estimates of the model

based equity risk premia of so-called "Small-Big", "Value-Growth" and "Winners-Losers" portfolios. Further, I find that they are all time-varying and all crashed in the 2008 financial crisis. Additionally, both the jump and volatility components of equity risk premia are especially important for the "Winners-Losers" portfolio.

To Jiyang Yu for his constant patience and support.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my advisor, Tim Bollerslev, and my committee members George Tauchen, Andrew Patton and Shakeeb Khan for their guidance and support. I also thank Hao Zhou, Dong Hwan Oh, Dan Lee, and the many financial econometric seminar participants at Duke University for many helpful discussions and suggestions.

1

Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence

1.1 Introduction

A number of recent studies have argued that aggregate U.S. stock market return is predictable over horizons ranging up to two quarters based on the difference between options-implied and actual realized variation measures, or the so-called variance risk premium (see, e.g., Bollerslev, Tauchen, and Zhou, 2009; Drechsler and Yaron, 2011a; Gabaix, 2012; Kelly, 2011; Zhou, 2010; Zhou and Zhu, 2013, among others). These findings are distinctly different from the longer-run multi-year return predictability patterns that have been studied extensively in the existing literature, in which the predictability is typically associated with more traditional valuation measures such as dividend yields, P/E ratios, or consumption-wealth ratios (see, e.g., Fama and French, 1988a; Campbell and Shiller, 1988b; Lettau and Ludvigson, 2001, among others). The present paper builds and further expands on the scope of these striking new empirical findings.

The variance risk premium is formally defined as the difference between the risk-neutral and statistical expectations of the future return variation.[1] It may be interpreted as a measure of both aggregate risk aversion and aggregate economic uncertainty. In our main empirical investigations reported on below, we follow Bollerslev, Tauchen, and Zhou (2009) in approximating the variance risk premium by the difference between one-month forward looking model-free options implied variances and the actual one-month realized variances at the time. This directly observable proxy has the obvious advantage of being simple to implement and completely model-free.

Our investigations are essentially threefold. First, to assess the validity of the previously documented predictability patterns, we report the results from a Monte Carlo simulation study designed to closely mimic the dynamic dependencies inherent in daily U.S. returns and variance risk premia. Our results clearly show that statistical biases can not "explain" the documented return predictability patterns.

Second, in a separate effort to corroborate and further expand on the existing empirical evidence based on monthly U.S. data prior to the advent of the financial crises, we extend the same basic return predictability regressions to seven other countries and more recent "out-of-sample" data spanning the financial crisis. We show that the same predictability pattern that exist for the U.S. hold true for most of the other countries, although the magnitudes are generally somewhat attenuated.

Third, motivated by this apparent commonality across countries, we define a "global" variance risk premium. We show that this simple aggregate world-wide variance risk premium results in strong predictability for all of the countries in the sample.

The finite sample properties of overlapping long-horizon return regressions have been studied extensively in the literature. Boudoukh, Richardson, and Whitelaw

[1] The variance risk premium is sometimes defined the other way around as the statistical minus risk-neutral expectations. This, of course, is immaterial for all of the results reported on below.

(2008), for instance, have recently shown that even in the absence of any increase in the true predictability, the values of the R^2's in regressions involving highly persistent predictor variables and overlapping returns will by construction increase roughly proportional to the return horizon and the length of the overlap.[2] By contrast, the variance risk premium is not especially persistent at the monthly horizon. Our simulations is based on a bivariate VAR-GARCH-DCC model designed to closely mimic the relevant joint dynamic dependencies in the daily return and variance risk premium. We find that the robust t-statistics usually employed in the literature are reasonably well behaved, albeit slightly over-sized under the null hypothesis of no predictability. We also find that the quantiles in the finite sample distribution of the R^2's from the regressions are spuriously increasing with the return horizon under the null of no predictability, which are distinctly different from the hump-shaped R^2s actually observed in the U.S. data at the 1-12 months horizons.

Guided by the Monte Carlo simulations, we rely on simple OLS regressions along with Newey-West based t-statistics based on simulated critical value to summarize our new international evidence. Due to data availability and liquidity considerations, we restrict our attention to the eight financial markets of France, Germany, Japan, Switzerland, the Netherlands, Belgium, the U.K., and the U.S. Regressing the individual country returns on the country specific variance risk premia result in similar hump-shaped regression coefficients and R^2's for all of the eight countries. However, the degree of predictability afforded by the country specific variance risk premia and the statistical significance of the regression coefficients are generally not as strong as the previously reported results for the U.S.

These results naturally raise the question of whether world-wide variance risk,

[2] Closely related issues pertaining to the use of persistent predictor variables have also been studied by, e.g., Stambaugh (1999), Ferson, Sarkissian, and Simin (2003), Baker, Taliaferro, and Wurgler (2006), Campbell and Yogo (2006), Ang and Bekaert (2007), and Goyal and Welch (2008), among others.

as opposed to the country specific variance risk, is being priced by the market? To investigate this idea, we construct a simple "global" variance risk premium proxy, defined as the market capitalization weighted average of the individual country variance risk premia. Restricting the effect on this "global" variance risk premium to be the same across countries in a panel return regression results in much stronger findings for all of the countries, with a systematic peak in the degree of predictability around the four month horizon. Moreover, the degree of predictability afforded by this "global" variance risk premium easily exceeds that of the implied and realized variation measures when included in isolation. It also clearly dominates that of other traditional predictor variables that have been shown to work well over longer annual horizons, including the P/E ratio.[3]

Our use of the variance difference as a simple proxy for the variance risk premium implicitly assumes that the volatility follows a random walk.[4] To investigate the sensitive of our main international findings to this simplifying assumption, we define a forward looking "global" variance risk premium from the differences between the individual countries one-month options implied variance and the corresponding one-month VAR-based forecasts for the actual variance. This alternative definition of the "global" variance risk premium give rise to almost identical international return predictability patterns.

Putting things into perspective, our new empirical findings are clearly related to the large existing literature on international stock return predictability (see, e.g.,

[3] Related evidence has also been reported in a few other recent studies pertaining to other markets. In particular, in concurrent independent work, Londono (2010) finds that the U.S. variance risk premium predicts several foreign stock market returns. In a slightly different context, Mueller, Vedolin, and Zhou (2011) argue that the U.S. variance risk premium predicts bond risk premia, beyond the predictability afforded by forward rates, while Buraschi, Trojani, and Vedolin (2010) and Zhou (2010) show that the variance risk premium also helps predict credit spreads, over and above the typical interest rate predictor variables.

[4] Of course, the variance difference may simply be interpreted as powerful predictor variable in its own right.

Harvey, 1991; Bekaert and Hodrick, 1992; Campbell and Hamao, 1992; Ferson and Harvey, 1993, among others). However, the focus of this literature has traditionally been on longer-run multi-year return predictability. By contrast, our results pertaining to the "global" variance risk premium concern much shorter-run within year predictability, and are essentially "orthogonal" to the findings reported in the existing literature.[5]

The rest of the paper is organized as follows. Section 2 presents our simulation-based results pertaining to the statistical inference procedures and robustness of the existing empirical evidence for the U.S. Section 3 details our international data and country specific return regressions. The results based on our new "global" variance risk premium and the combined panel regressions for all of the countries, are discussed in Section 4. Section 5 concludes.

1.2 General Setup and Monte Carlo Simulations

The key empirical findings reported in Bollerslev, Tauchen, and Zhou (2009) (BTZ, henceforth), and the subsequent studies cited above, are based on simple OLS regressions of the returns on the aggregate market portfolio over monthly and longer return horizons on a measure of the one-month variance risk premium.

In particular, let $r_{t,t+\tau}$ and VRP_t denote the continuously compounded return from time t to time $t+\tau$ and the variance risk premium at time t, respectively. Defining the unit time interval to be one month, the multi-period return regressions in BTZ may then be expressed as,

$$\frac{1}{h}\sum_{j=1}^{h} r_{t+(j-1),t+j} \;=\; a(h) \;+\; b(h)VRP_t \;+\; u_{t,t+h} \tag{1.1}$$

[5] Other recent studies highlighting short-run international predictability include Rapach, Strauss, and Zhou (2010) based on lagged U.S. returns, Ang and Bekaert (2007) and Hjalmarsson (2010) based on short-term interest rates, and Bakshi, Panayotov, and Skoulakis (2011) based on the Baltic Dry Index.

where $t = 1, 2, ..., T - h$ refer to the specific observations used in the regression.

Meanwhile, it is well known that in the context of overlapping return observations, the regression in (1) can result in spuriously large and highly misleading regression R^2's, say $R^2(h)$, as the horizon h increases; see, e.g., the discussion and many references in Campbell, Lo, and MacKinlay (1997). Similarly, the standard errors for the OLS estimates designed to take account of the serial correlation in $u_{t+h,t}$ based on the Bartlett kernel advocated by Newey and West (1987) (NW, henceforth), and the modification proposed by Hodrick (1992) (HD, henceforth), can also both result in t-statistics for testing hypotheses about $a(h)$ and $b(h)$ that are poorly approximated by a standard normal distribution.

Most of the existing analyses pertaining to these and other related finite sample biases, however, have been calibrated to situations with a highly persistent predictor variable, as traditionally used in long-horizon return regressions. Even though the variance risk premium is fairly persistent at the daily frequency, it is much less so at the monthly level, and as such one might naturally expect the finite sample biases to be less severe in this situation.[6] Our Monte Carlo simulations discussed in the next section confirm this conjecture in an empirically realistic setting designed to closely mimic the joint dependencies in actual daily returns and variance risk premia.

1.2.1 Simulation Design

The model underlying our simulations is based on daily S&P 500 composite index returns (obtained from CRSP). The corresponding daily observations on the variance risk premium are defined as $VRP_t = IV_t - RV_{t-1,t}$, where we rely on the square of the new VIX index (obtained from the CBOE) to quantify the implied variation IV_t,

[6] The first order autocorrelation coefficient for the monthly U.S. variance risk premium analyzed in the empirical section below equals 0.39, and it is even lower for all of the other countries included in our subsequent analysis. By comparison, the first order autocorrelations for monthly dividend yields, P/E ratios, and other valuation ratios typically employed in the long-horizon regression literature, are around 0.95-0.99.

and the summation of current and previous 20 trading days daily realized variances (obtained from Standard & Poor's) together with the squared overnight returns to quantify the total realized variation over the previous month $RV_{t-1,t}$.[7]

The sample period runs from February 1, 1996 to December 31, 2007, for a total of 2,954 daily observations. The end of the sample purposely coincides with that in BTZ. We will later investigate the sensitivity of the empirical results to the inclusion of more recent data involving the financial crisis. The span of the data exactly matches the length of the commonly available sample for the eight countries analysis below.

After some experimentation, we arrived at the following bivariate VAR(1)-GARCH(1,1)-DCC model (see Engle, 2002, for additional details on the DCC model) for the two daily time series, define $\Delta = 1/20$,

$$
\begin{aligned}
r_{t-\Delta,t} &= \underset{(0.001)}{-1.958\text{e-}5} - \underset{(0.016)}{0.009} r_{t-2\Delta,t-\Delta} + \underset{(0.010)}{0.025} VRP_{t-\Delta} + \epsilon_{t,r} \\
VRP_t &= \underset{(0.001)}{3.759\text{e-}5} + \underset{(0.017)}{0.033} r_{t-2\Delta,t-\Delta} + \underset{(0.010)}{0.972} VRP_{t-\Delta} + \epsilon_{t,vrp} \\
\sigma^2_{t,r} &= \underset{(1.68\text{e-}6)}{1.280\text{e-}6} + \underset{(0.004)}{0.071} \epsilon^2_{t-\Delta,r} + \underset{(0.008)}{0.920} \sigma^2_{t-\Delta,r} \\
\sigma^2_{t,vrp} &= \underset{(7.59\text{e-}6)}{2.038\text{e-}7} + \underset{(0.004)}{0.133} \epsilon^2_{t-\Delta,vrp} + \underset{(0.028)}{0.871} \sigma^2_{t-\Delta,vrp} \\
Q_t &= \begin{pmatrix} \underset{(0.036)}{0.997} & \underset{(0.040)}{-0.754} \\ \underset{(0.040)}{-0.754} & \underset{(0.060)}{1.023} \end{pmatrix} + \underset{(0.002)}{0.011} \eta_{t-\Delta} \eta'_{t-\Delta} + \underset{(0.004)}{0.979} Q_{t-\Delta} \\
R_t &= diag\{Q_t\}^{-1} Q_t diag\{Q_t\}^{-1},
\end{aligned}
$$

where $\eta_t \equiv (\frac{\epsilon_{t,r}}{\sigma_{t,r}}, \frac{\epsilon_{t,vrp}}{\sigma_{t,vrp}})'$, and $E_{t-\Delta}(\eta_t) = 0$ and $E_{t-\Delta}(\eta_t \eta'_t) = R_t$ by assumption. The specific parameter values refer to Quasi Maximum Likelihood Estimates (QMLE)

[7] This directly mirrors the definition of the variance risk premium employed in BTZ. Forward looking measures of VRP_t that align IV_t with a measure of the expected volatility $E_t(RV_{t,t+1})$ have also been used in the literature. However, this requires additional modeling assumptions for calculating $E_t(RV_{t,t+1})$, whereas the VRP_t used here has the obvious advantage of being directly observable at time t. We will return to this issue in Section 4 below.

obtained under the auxiliary assumption of conditional normality, with robust standard errors following Bollerslev and Wooldridge (1992) in parentheses. With the exception of the lagged daily returns, most of the dynamic coefficients are highly significant at conventional levels.

The model implies a strong negative (on average) correlation between the innovations to the return and VRP equations. This, of course, is consistent with the well documented "leverage" effect; see, e.g., Bollerslev, Sizova, and Tauchen (2012b) and the many references therein. At the same time, as is evident from the equation for Q_t, the conditional correlation clearly varies over time, and as shown in the top panel in Figure 1.1 reaches a low of close to -0.85 toward the end of the sample. The bottom three panels in Figure 1.1 indicate that the distribution of the estimated standardized residuals from the model (i.e., $\widehat{c\eta}_t \equiv \widehat{F}_t^{-1}\widehat{\eta}_t$, where $\widehat{F}_t \cdot \widehat{F}_t' = \widehat{R}_t$) are well behaved and centered at zero, with variances close to unity, albeit not normally distributed.[8] Thus, all in all the model provides a reasonably good fit to the joint dynamic dependencies inherent in the two daily series.

As such, we will use this relatively simple-to-implement model as our basic data generating process for the Monte Carlo simulations, and our analysis of the finite sample properties of the NW and HD t-statistics, and $R^2(h)$'s from the overlapping return regressions in equation (1).[9] Our simulated finite sample distributions will be based on a total of 2,000 bootstrapped replications from the model. We will look at monthly sample frequencies, and return horizons h ranging up to 12 months. The

[8] The sample means for $\widehat{c\eta}_{t,1}$ and $\widehat{c\eta}_{t,2}$ equal -0.044 and 0.088, the standard deviations equal 0.999 and 1.007, while the skewness and kurtosis equal -0.469 and 0.894, and 4.913 and 7.860, respectively. Further diagnostic checks also reveal that while the residuals from the return equation appear close to serially uncorrelated, there is some evidence for neglected longer-run serial dependencies in the equation for the variance risk premium.

[9] The bandwidth in the Bartlett kernel employed in our implementation of the NW standard errors is set to $m = [h + 4 * ((T - hs)/100)^{2/9}]$, where $[\cdot]$ refers to the integer value. We also experimented with the reverse regression technique suggested by Hodrick (1992) for testing $b^s(h) = 0$. The results, available upon request, were very similar to the ones for the HD t-statistic reported below.

number of observations for each of the simulated samples is fixed at 149 months, corresponding to the length of the actual sample used in the estimation of the VAR-GARCH-DCC model above.[10] We begin with a discussion of the size and power properties of the two t-statistics.

1.2.2 Size and Power

Our characterization of the distributions under the null hypothesis of no return predictability is based on restricting the coefficients associated with $r_{t-2\Delta,t-\Delta}$ and $VRP_{t-\Delta}$ in the return equation to be identically equal to zero, leaving all of the other coefficients at their estimated values. In Table 1.1, panel A reports the resulting simulated 95th percentiles of the t^{NW} and t^{HD} test statistics, along with the regression R^2's. In line with the evidence in the existing literature, both of the t-statistics exhibit non-trivial size distortions relative to the nominal one-side 95-percent critical value of 1.645. Also, the distortions tend to increase with the return horizon h. Moreover, consistent with the results reported in Hodrick (1992), the biases for the NW based standard error calculations generally exceed those for the HD standard errors, and markedly more so the longer the return horizon.

To illustrate the results, we plot in the top left panel in Figure 1.2 the simulated 95-percent critical values for the t^{NW} (dashed lines) and the t^{HD} (solid lines) statistics for monthly sampled data. We also include in the figure the t-statistics obtained by running these same regressions on the monthly data over the February 1996 through December 2007 sample period used in calibrating the simulated model. As the figure shows, the actual t^{NW}-statistics exceeds the simulated critical values for return horizons in the range of 2 to 3 months. Meanwhile, the t^{HD}-statistics generally do not exceed the simulated critical values and accordingly do not support the idea

[10] As previously noted, this also mimics the length of the commonly available sample for the international data analyzed below.

of return predictability.

In order to better understand this discrepancy in the conclusions drawn from the two tests, we report panel B in Table 1.1 the power of the tests to detect predictability implied by the unrestricted VAR-GARCH-DCC model. To facilitate comparisons we only report the size-adjusted power for a 5-percent test. Not surprisingly, the power of both tests decrease with the return horizon. But, the power of the t^{NW} test exceed that of the t^{HD} test for return horizons less than a year, and the differences appear most pronounced at the 1-4 month horizons. These differences are also evident in the top right panel in Figure 1.2, which plot the relevant power curves.

In addition to the t-statistics associated with the $b(h)$ coefficients, the $R^2(h)$'s from the return regressions are also commonly used to assess the strength of the relationship and the effectiveness of the predictor variable across different horizons. Of course, it is well known that the biases exhibited by the t-statistics in the context of long-horizon return regressions with persistent predictor variable carry over to the $R^2(h)$'s, and that these need to be carefully interpreted as well (see, e.g., the aforementioned study by Boudoukh, Richardson, and Whitelaw, 2008, for a recent analysis, along with the many references therein).

The corresponding columns in Table 1.1 show that, while less dramatic than the biases that exist over multi-year return horizons with highly persistent predictor variables, the $R^2(h)$'s can still be quite different from zero under the null of no predictability in the present setting. In particular, the 95th percentiles are around 5-6 percent at the 2-4 months horizon.

Further to this effect, we show in the bottom left panel in Figure 1.2 select quantiles in the simulated distribution of the $R^2(h)$'s from daily regression that obtain in the absence of any predictability. Consistent with the findings in the extant literature pertaining to monthly observations and longer return horizons, all of the quantiles increase monotonically with the return horizon, and this increase is

especially marked for the higher percentiles. Intuitively as the horizon increases, the overlapping return regressions become closer to a spurious type regression.

In addition to the simulated quantiles, the bottom right panel of the figure also shows the $R^2(h)$'s obtained from the monthly return regressions implied by the same VAR-GARCH-DCC model. Comparing the actual $R^2(h)$'s to the simulated percentiles again suggest that the degree of predictability is most significant at the intermediate 2-4 months horizon. This, of course, is directly in line with the inference based on the t-statistics discussed in the previous section. It also supports the prior empirical evidence reported in BTZ.

The hump-shaped pattern in the actual $R^2(h)$'s, with an apparent peak in the degree of predictability at the intermediate 2-4 months horizon, also closely mimics the patterns in the simulated quantiles for the estimated VAR-GARCH-DCC model depicted in the bottom panel in Figure 1.2. Interestingly, this striking similarity arises in spite of the fact that the simulated model involves only first-order dynamics in the equations that describe the daily conditional means.

Taken as a whole, our Monte Carlo simulations and the new regression results based on daily U.S. returns discussed above clearly support the variance risk premium as a powerful predictor at the 2-4 month horizons. At the same time, the overlapping nature of the return regressions tend to attenuate the strength of the predictability somewhat. Hence, in an effort to further corroborate the existing empirical evidence pertaining exclusively to the U.S. market and data prior to the 2008 financial crisis, we next turn to a discussion of our new empirical findings involving more recent data and several other countries. For each country considered, we will base our empirical investigations on monthly predictive regression and NW-based standard errors with simulated critical values.

1.3 International Evidence

Motivated by the Monte Carlo simulation results, we will rely on the common benchmark monthly OLS regressions, along with the simulated NW critical values and t^{NW}-statistics for characterizing the return predictability internationally, keeping in mind the finite sample biases documented in the simulations. We will restrict our analysis to France, Germany, Japan, Switzerland, the Netherlands, Belgium, the U.K., and the U.S., all of which have highly liquid options markets and readily available model-free implied variances for their respective aggregate market indexes (see Siriopoulos and Fassas, 2009, for a recent summary of the model-free and parametric options implied volatility indexes available for different countries). We begin with a brief discussion of the data.

1.3.1 Data and Summary Statistics

Our monthly aggregate market returns for the different countries are based on data for the French CAC 40 (obtained from Euronext), the German DAX 30 (obtained from Deutsche Börse), the Japanese Nikkei 225, the Swiss SMI, the Netherlands AEX, the Belgium BEL 20, and the U.K. FTSE 100 (all obtained from Datastream), and the U.S. S&P 500 (obtained from Standard & Poor's). We use the sum of the daily squared returns over a month to construct end-of-month realized variances RV_t^i for each of the countries. We obtain the corresponding end-of-month model-free implied volatilities $(IV_t^i)^{1/2}$ for the S&P 500 (VIX) from the CBOE, the CAC (VCAC) from Euronext, the DAX (VDAX) from Deutsche Börse, while those for the FTSE (VFTSE), the SMI (VSMI), AEX (VAEX) and BEL (VBEL) were obtained from Datastream. Our data for the Japanese volatility index (VXJ) was obtained directly from the Center for the Study of Finance and Insurance at Osaka University (see Nishina, Maghrebi, and Kim, 2006, for a more detailed discussion of the VXJ

index). Finally, the risk-free rates used in the construction of the excess returns were obtained from the Federal Reserve Board and Eurocurrency via Datastream.[11]

The sample period for each of the series extends from January 2000 to December 2011. The beginning of the sample coincides with the back-dated initial date of the NYSE Euronext volatility indices.[12] The use of more recent data through 2011 allows for additional validation of the original empirical evidence for the U.S. based on data prior to the financial crisis.

In accordance with the empirical analysis in the previous section, the proxy for the variance risk premium for each of the individual countries is simply defined by $VRP_t^i \equiv IV_t^i - RV_{t-1,t}^i$. As noted above, this proxy has the obvious advantage of being directly observable. The time series plots of VRP_t^i for each of the eight countries in Figure 1.3 clearly show the dramatic impact of the financial crisis, and the exceptionally large variance risk premia observed in the Fall of 2008. Interestingly, however, the premium for the DAX, and to a lesser extent the SMI, were almost as large and negative in 2001-2002.

The standard set of summary statistics reported in Table 3.1 also show a remarkable coherence in the distributions of the variance risk premia and monthly excess returns across countries. In particular, looking at Panel A, the average excess returns all reflect the often-called "lost decade," ranging from a high of -2.54 for Switzerland to a low of -9.26 for Belgium. Of course, the corresponding standard deviations all point to considerable variations in the returns around their negative sample means.

The variance risk premia are almost all positive on average, ranging from a low of -2.74 for Belgium to a high of 12.32 for Japan on a percentage-squared monthly

[11] The use of excess returns, as opposed to raw returns, has almost no effect on the results from the return predictability regressions reported below.

[12] The volatility indexes is available prior to January 2000 for some of the countries; VDAX (December 1994), VXJ (January 1998), VSMI (January 1999) and VIX (January 1990). Comparable results to the ones for the country specific regressions discussed below based on the longest possible sample for each of the countries are reported in a Supplementary Appendix available upon request.

basis. "Selling" volatility has been highly profitable on average over the last decade. Meanwhile, consistent with the visual impressions from Figure 1.3, all of the premia are significantly negatively skewed and exhibit large excess kurtosis. Even though the implied and realized variances are both strongly serially correlated for all of the countries, the variance risk premia are generally not very persistent, and the maximum first order serial correlation observed for the S&P 500 equals just 0.39. Turning to Panels B and C, the sample cross-country correlations are all fairly high, and with the exceptions of those for the Nikkei and Belgium, the correlations for the returns all exceed 0.80, while those for the variance risk premia are in excess of 0.70.

The similarities in the summary statistics in Table 3.1 and the time series plots in Figure 1.3, naturally suggest that the same predictive relationship documented for the U.S. returns and variance risk premium may hold true for the other countries. The results discussed in the next subsection generally corroborate this conjecture.

1.3.2 Country Specific Regressions

In parallel to the general multi-period return regressions defined in (1.1), our monthly return regressions for each of the individual countries may be conveniently expressed as,

$$h^{-1}r^i_{t,t+h} \;=\; a^i(h) \;+\; b^i(h)VRP^i_t \;+\; u^i_{t,t+h}\;, \tag{1.2}$$

where $r^i_{t,t+h}$ and VRP^i_t refer to the $h = 1, 2, ..., 12$ month excess return and variance risk premium for country i, respectively.

The actual estimates for $b^i(h)$ and the corresponding t^{NW}-statistics reported in Table 2.4 obviously differ somewhat across countries. However, with the exception of France, Belgium and the U.S., the estimated coefficients all show the same general pattern starting out fairly low and insignificant at the shortest one-month horizon, rising to their largest values at 3-5 months, and then gradually tapering off thereafter for longer return horizons. These similarities are also evident in Figure 1.4, which

displays the regression coefficients along with their 90% NW standard error bands according to our simulated critical value in the simulation section.[13]

These similarities in the patterns in the estimated $b(h)$ coefficients naturally translate into very similar patterns in the regression $R^2(h)$'s as well. In particular, looking at the plots in Figure 1.5, all of the $R^2(h)$'s exhibit an almost identical hump-shaped pattern with the degree of predictability maximized around the 4 months horizon. Of course, the actual values of the $R^2(h)$'s vary somewhat across the different country indices, achieving a maximum of only 0.60 percent for the Nikkei 225 compared to 13.03 percent for the S&P 500.[14] Interestingly, this value of $R^2(3) = 13.02$ for the U.S. exceeds that obtained with monthly data through the end of 2007 previously reported in BTZ and Drechsler and Yaron (2011a).

The qualitative results from the country specific VRP regressions, while not as significant, are generally in line with the existing results for the U.S. Going one step further, the similarities in the patterns observed across the different countries also suggest that even stronger results may be available by pooling the regressions and entertaining the notion of a common "global" variance risk premium. We explore these ideas next.

[13] We also try Stambaugh correction for the country specific regression, and we find that the estimated bias is negligible. In fact, variance risk premiums at monthly frequency are much less persistent, and the contemporaneous correlation between residuals of a bivariate VAR are only slightly negatively correlated. The Stambaugh correction results are reported in the Supplementary Appendix.

[14] This lack of predictability for Japan is also consistent with the evidence reported in Ubukata and Watanabe (2011).

1.4 Global Variance Risk

Our proxy for the "global" variance risk premium is based on a simple capitalization weighted average of the proxies for country specific variance risk premia,

$$VRP_t^{global} \equiv \sum_{i=1}^{8} w_t^i VRP_t^i ,$$

where $i = 1, 2, ..., 8$ refer to each of the eight countries included in our analysis.[15] The end-of-month market capitalizations used in defining the weights w_t^i are obtained from Thomson Reuters Institutional Brokers' Estimate System (I/B/E/S) via Datastream. The plot of the weights in Figure 1.6 shows that the U.S. market accounts for around sixty percent through most of the sample period, with the Japanese market a distant second. This large weight assigned to the U.S. market in our definition of the "global" VRP index is also implicit in the aforementioned summary statistics in Panel C in Table 3.1, and the relatively high correlation of 0.89 between VRP_t^{global} and $VRP_t^{S\&P500}$.

1.4.1 Individual Country Regressions

The results for the regressions obtained by replacing the country specific VRP_t^i's in equation (2) with the new VRP_t^{global} proxy,

$$h^{-1} r_{t,t+h}^i = a^i(h) + b^i(h) VRP_t^{global} + u_{t,t+h}^i , \tag{1.3}$$

are reported in Table A. Comparing the results to the ones for the country specific regressions in Table 2.4, reveals even stronger commonalities and uniform patterns across countries. The "global" VRP proxy serves as a highly significant predictor variable for all of the different country returns, with t^{NW}-statistics systematically

[15] This parallels the construction used in Harvey (1991) in the estimation of the world price of covariance risk.

in excess of 4.0 at the 4 or 5 months horizon. Further increasing the horizon h, VRP_t^{global} systematically becomes insignificant for predicting the longer 9 and 12 months returns.

These striking cross-country similarities are also evident from the plots of the estimated regression coefficients and the 90% NW-based confidence bands with simulated critical values in Figure 1.7. Not only do the individual country estimates for the $b^i(h)$'s look very similar, the confidence bands also tend to be tighter compared to the country specific regressions discussed above. Further along these lines, Figure 1.8 shows the general patterns in the predictability, as measured by the $R^2(h)$'s, to be very similarly shaped across countries, with peaks at the 4-5 months return horizon.[16]

These remarkable similarities in the estimates for the different countries, naturally suggest restricting the coefficients in equation (1.3) to be the same across countries, as a way to enhance the efficiency of the estimates and to ensure a common reward for bearing "global" variance risk.

1.4.2 Panel Regressions

The estimation results from the panel regression that restricts the coefficients for the "global" variance risk premium to be the same across countries,

$$h^{-1}r^i_{t,t+h} \;=\; a(h) \;+\; b(h)VRP_t^{global} \;+\; u^i_{t,t+h}\;, \tag{1.4}$$

are reported in Table A (for additional details on calculating standard errors, see, e.g., Petersen, 2009).[17] As the table clearly shows, the use of panel regressions do

[16] The relatively large weight assigned to the U.S. in our construction of the "global" variance risk premium means that fairly similar results are obtained by replacing the new VRP_t^{global} in the regressions in equation 1.3 with $VRP_t^{S\&P500}$. These additional results are available upon request. Comparable empirical results based on the U.S. variance risk premium have also recently been reported in concurrent independent work by Londono (2010), who ascribes the predictability to informational frictions along the lines of Rapach, Strauss, and Zhou (2010).

[17] We also experimented with the two-way cluster analysis in Cameron, Gelbach, and Miller (2011), resulting in very similar findings.

indeed result in more accurate estimates, and a highly significant t^{NW}-statistics of 10.91 at the 4-months horizon. The average panel regression $R^2(h)$'s for the eight countries also gradually rise from less than 2 percent at the one-month horizon to a large 6.21 percent for the 4-month returns, tapering off to zero for the longer 9-12 month return horizons.

These key empirical findings are succinctly summarized in Figure A.2, which plots the panel regression estimates for the $b(h)$s based on the country specific VRP's and the "global" VRP proxy along with their two NW-based standard error bands (top two panels), and the corresponding panel regression $R^2(h)$'s (bottom two panels). The VRP^{global}-based regressions (depicted in the right two panels) obviously result in sharper coefficient estimates and stronger average predictability across the eight countries than do the individual country VRP^i regression (depicted in the left two panels).

The average panel regression $R^2(h)$'s, of course, mask important cross-country differences in the degree of predictability. We therefore also show in Figure A.3 the country specific implied $R^2(h)$'s obtained by evaluating the individual country regressions in equation (1.3) at the more precisely estimated common $\hat{a}(h)$ and $\hat{b}(h)$ obtain from the panel regressions in equation (1.4). Comparing Figure A.3 to the earlier Figure 1.8 for the individual country regressions, it is clear that the added precision afforded by restricting the $a^i(h)$ and $b^i(h)$ coefficients to be the same across countries sacrifices very little in terms of the implied predictability.

1.4.3 Robustness Checks

To assess the robustness of these striking international predictability patterns, the next panel in Table A reports the results obtained by including a capitalization weighted average of the country specific P/E ratios as an additional regressor in equation (1.4). Consistent with the results for the U.S. market in isolation reported in

BTZ, the "global" P/E ratio adds nothing to the predictability afforded by VRP^{global} within the one-year horizons reported in the table, leaving all of the estimates for $b(h)$ and the $R^2(h)$'s almost the same. The predictability of the "global" variance risk premium is effectively "orthogonal" to that documented in the existing literature based on more traditional macro-finance variables, such as the P/E ratio, dividend yields, and consumption-wealth ratios, which are typically only significant over longer multi-year return horizons (see, e.g., the classic studies by Fama and French, 1988a; Campbell and Shiller, 1988b; Lettau and Ludvigson, 2001).[18]

To further highlight the predictive gains afforded by the use of our "global" VRP as opposed to the own country VRP's, the last two panels in Table A show the estimates obtained by including each individual country's premium in a panel regression in place of VRP^{global},

$$h^{-1}r^i_{t,t+h} \;=\; a(h) \;+\; b(h)VRP^i_t \;+\; u^i_{t,t+h}\;. \tag{1.5}$$

While the results still point to overall efficiency gains from the use of the panel regression relative to the country specific regressions in Table 2.4, the magnitude of the return predictability is obviously much lower than for VRP^{global}. The "global" variance risk premium is clearly a much better predictor of the future returns for most of the countries than the individual country specific premia. Again, including the country specific P/E ratios in the same panel regression do not material affect the overall predictability as measured by the R^2s, nor the values of the estimated regression coefficients for the variance risk premia.

[18] Further corroborating the results for the U.S. market in BTZ, we also found that including the implied "global" variance or the realized "global" variance together with the "global" variance risk premium resulted in mostly insignificant coefficient estimates. These additional results are available upon request.

1.4.4 Forward Looking Global Variance Risk Premium

Our proxy for the "global" variance risk premium underlying our main findings discussed above is based on a weighted average of the variance difference for each the countries. This directly mirrors the original proxy for the U.S. variance risk premium employed in BTZ, and the proxy used in the country specific regressions in Section 3. To assess the sensitivity of our results to this simple and easy-to-implement proxy, we briefly summarize the results obtained by replacing the model-free lagged monthly realized variances with forward looking model-based expectations in the way we define the "global" variance risk premium.

Specifically, let $E_t(RV^i_{t,t+1})$ denote the time t expectation of the one-month ahead return variation for country i. Additionally, let $FVRP^i_t = IV^i_t - E_t(RV^i_{t,t+1})$ denote the corresponding forward looking variance risk premia for country i. We then define a forward looking "global" variance risk premium by,

$$FVRP^{global}_t \equiv \sum_{i=1}^{8} w^i_t FVRP^i_t \ .$$

In contrast to the VRP^{global} defined above, $FVRP^{global}$ necessitates the use of a model for generating the forward expectations $E_t(RV^i_{t,t+1})$. In the results reported on below, we follow Andersen, Bollerslev, and Diebold (2007) and Corsi (2009a) in generating these forecasts from HAR-RV type models in which we regress $RV^i_{t,t+1}$ for each of the eight countries on the daily, weekly, and monthly realized variances, $RV^i_{t-\Delta,t}$, $RV^i_{t-5\Delta,t}$, and $RV^i_{t-1,t}$ (where $\Delta = 1/20$), respectively, along with the options implied variances, IV^i_t, for all of the other seven countries.[19]

[19] We make sure that all of the regressors are properly aligned to correct for the different time zones, so that none of the predictions involve any future information. We also experimented with the use of a standard VAR(1) model involving only the current monthly realized variation measures, $RV^i_{t-1,t}$, and options implied variation measures, IV^i_t, for generating $E_t(RV^i_{t,t+1})$, resulting in qualitatively similar, albeit not as significant, predictive return regressions. Further details concerning these additional results are summarized in the Supplementary Appendix available upon request.

The resulting $FVRP^{global}$ is plotted in Figure A.4 (bottom panel), together with the previously used simple VRP^{global} proxy (top panel). While the two series obviously differ, the general dynamic dependencies are obviously quite similar. The large negative spike in VRP^{global} observed at the height of the financial crisis is slightly diminished in the forward looking $FVRP^{global}$.

Turning to the predictive return regressions, the top panel in Table 1.6 reports the estimates from the same panel regressions in equation (1.4) using $FVRP^{global}$ in place of VRP^{global}. While the NW-based t-statistics for the 1-6 month returns are all slightly lower than the comparable t^{NW}-statistics reported in the top panel in Table A, they remain highly significant at any reasonable level. In fact, the statistical significance of the regressions based on $FVRP^{global}$ extends to at least the 9 month horizon. The R^2s also show a similar hump shaped pattern to the ones in Table A and Figure A.2, with the predictability now maximized at the slightly longer 5-6 month horizon. This shift in the location of the peak is also consistent with the Monte Carlo results in Figure 1.2, and the slightly smaller first order autocorrelation of 0.31 for $FVRP^{global}$ compared to 0.36 for VRP^{global}. The second panel in Table 1.6 again further corroborates our key empirical findings, and the idea that the predictability inherent in the "global" variance risk premium is essentially "orthogonal" to that in the "global" P/E ratio, which only kicks in over longer annual horizons.

In sum, the estimated regression coefficients for the "global" variance risk premium are fairly similar across countries, and with the exception of the U.S., the $R^2(h)$'s for the panel regressions are generally larger for the "global" VRP than for the "local" VRP's.

These empirical findings are directly in line with a stylized two-country consumption based equilibrium model. We show that the "global" variance risk premium's includes relatively larger amount of the aggregate volatility uncertainty than the local variance risk premium from the smaller country—of which the consumption weight

is less than one half, directly mirroring the "global" variance risk premium's stronger return predictability. Conversely, for the larger country (the U.S.), the "local" VRP gives rise to marginally higher slope coefficients than the "global" VRP.[20]

1.5 Conclusion

A number of recent studies have argued that aggregate U.S. stock market return is predictable over relatively short 3-5 month horizons by the difference between options implied and actual realized variation, or the so-called variance risk premium. We show that this newly documented predictability is not due to finite sample biases in the statistical inference procedures, and that the apparent hump-shaped pattern in the degree of predictability documented in over-lapping monthly returns regressions is entirely consistent with the implications from an empirically realistic bivariate daily time series model for the returns and variance risk premia.

Further corroborating the existing empirical evidence for the U.S., we show that the same basic predictive relationship between future returns and current variance risk premia holds true with more recent "out-of-sample" data through 2011. We also show that the same basic results hold true for a set of seven other countries, although the magnitude of the predictability and the statistical significance of the own country variance risk premium tend to be somewhat muted relative to those observed for the U.S.

Meanwhile, employing a capitalization weighted "global" variance risk premium results in very similar shaped predictability patters across return horizons for *all* of the countries in our sample, and uniformly larger t-statistics. Further restricting the regression coefficients and the compensation for "global" variance risk to be the same across countries, we find even stronger results and highly significant test statistics,

[20] Details of the model setup and calibration evidence can be found in the online Supplementary Appendix.

with the degree of predictability maximized at the 4-5 month horizon.

The "global" variance risk premium may be seen as a proxy for world-wide aggregate economic uncertainty, therefore "global" variance risk premium generally provides more accurate predictions of the future individual country returns than the own country variance risk premia. Alternatively, the "global" variance risk premium may be interpreted as a measure of aggregate risk aversion (e.g., Bekaert, Engstrom, and Xing, 2009), or a summary measure of world-wide disagreements in beliefs (e.g., Buraschi, Trojani, and Vedolin, 2010). All of these different economic mechanisms likely play some role in generating the strong international return predictability embodied in the "global" variance risk premium first documented here. We leave it for future research to sort out this important question.

Table 1.1: Simulated Size, Power and R^2

The top panel reports the simulated 95-percentiles in the finite sample distributions of t^{NW} and t^{HD} for testing the hypothesis that $b_s(h) = 0$ based on the return predictability regression in equation (1), along with the adjusted R^2 from the regression. The data are generated from the VAR-GARCH-DCC model discussed in the main text, restricting the coefficients in the conditional mean equation for the returns to be equal to zero. The lower panel reports the simulated power of the size-adjusted 5-percent t^{NW} and t^{HD} statistics for testing the null hypothesis of no predictability and $b_s(h) = 0$ in the return regression in equation (1). The data are generated from the VAR-GARCH-DCC model discussed in the main text. In both size and power studies, the "monthly" regressions involve 149 observations, and the simulations are based on a total of 2,000 replications.

Panel A: Simulated Size and R^2

Horizon	1	2	3	4	5	6	9	12
t^{NW}	2.2602	2.5199	2.7876	2.9413	3.2413	3.2200	3.3143	3.5087
t^{HD}	2.2763	2.1871	2.0835	2.1063	2.1024	2.1237	2.1631	2.1857
$adj.R^2$	3.0169	4.8366	5.7740	6.3148	7.4592	7.5017	8.1923	8.6792

Panel B: Simulated Power

Horizon	1	2	3	4	5	6	9	12
pw^{NW}	0.8865	0.8450	0.7680	0.6855	0.5625	0.5070	0.3680	0.2770
pw^{HD}	0.8070	0.7625	0.7105	0.6265	0.5470	0.4970	0.3500	0.3025

Table 1.2: Summary Statistics

The monthly excess returns are in annualized percentage form. The variance risk premia are in monthly percentage-squared form. The "global" index of variance risk premium are defined in the main text. The sample period extends from January 2000 to December 2011.

Panel A: Excess Returns and Variance Risk Premia

	AEX		BEL 20		CAC 40		DAX 30		FTSE 100		Nikkei 225		SMI		S&P 500		Global Index
	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	VRP_t
Mean	-9.26	6.37	-5.25	-2.74	-8.64	2.75	-5.00	4.68	-4.93	7.55	-7.59	12.32	-2.54	6.45	-2.94	7.40	7.38
Std. Dev	76.59	43.30	65.02	43.28	67.49	41.69	82.81	33.43	52.35	32.16	71.73	38.47	51.90	28.10	57.52	35.47	32.74
Skewness	-0.99	-4.47	-1.36	-4.58	-0.59	-4.84	-0.90	-2.83	-0.64	-5.56	-0.76	-4.88	-0.70	-3.25	-0.61	-4.71	-5.46
Kurtosis	5.01	35.00	6.44	32.81	3.60	41.98	5.49	15.98	3.56	50.72	4.66	42.71	3.50	23.86	3.92	35.55	47.52
AR(1)	0.12	0.35	0.30	0.31	0.13	0.30	0.09	0.10	0.07	0.34	0.13	0.16	0.27	0.14	0.15	0.39	0.36

Panel B: Correlation Matrix for Excess Returns

	AEX	BEL 20	CAC 40	DAX	FTSE 100	Nikkei 225	SMI	S&P 500
AEX	1.00	0.86	0.91	0.88	0.87	0.61	0.83	0.81
BEL 20		1.00	0.83	0.76	0.81	0.52	0.79	0.75
CAC 40			1.00	0.93	0.90	0.60	0.84	0.87
DAX				1.00	0.85	0.57	0.80	0.83
FTSE 100					1.00	0.62	0.80	0.88
Nikkei 225						1.00	0.58	0.64
SMI							1.00	0.78
S&P 500								1.00

Panel C: Correlation Matrix for Variance Risk Premia

	AEX	BEL 20	CAC 40	DAX	FTSE 100	Nikkei 225	SMI	S&P 500	Global
AEX	1.00	0.85	0.91	0.86	0.92	0.64	0.86	0.81	0.85
BEL 20		1.00	0.81	0.68	0.80	0.51	0.77	0.69	0.93
CAC 40			1.00	0.84	0.89	0.65	0.79	0.82	0.86
DAX				1.00	0.78	0.54	0.86	0.70	0.92
FTSE 100					1.00	0.73	0.84	0.88	0.64
Nikkei 225						1.00	0.60	0.64	0.87
SMI							1.00	0.69	0.81
S&P 500								1.00	0.89
Global									1.00

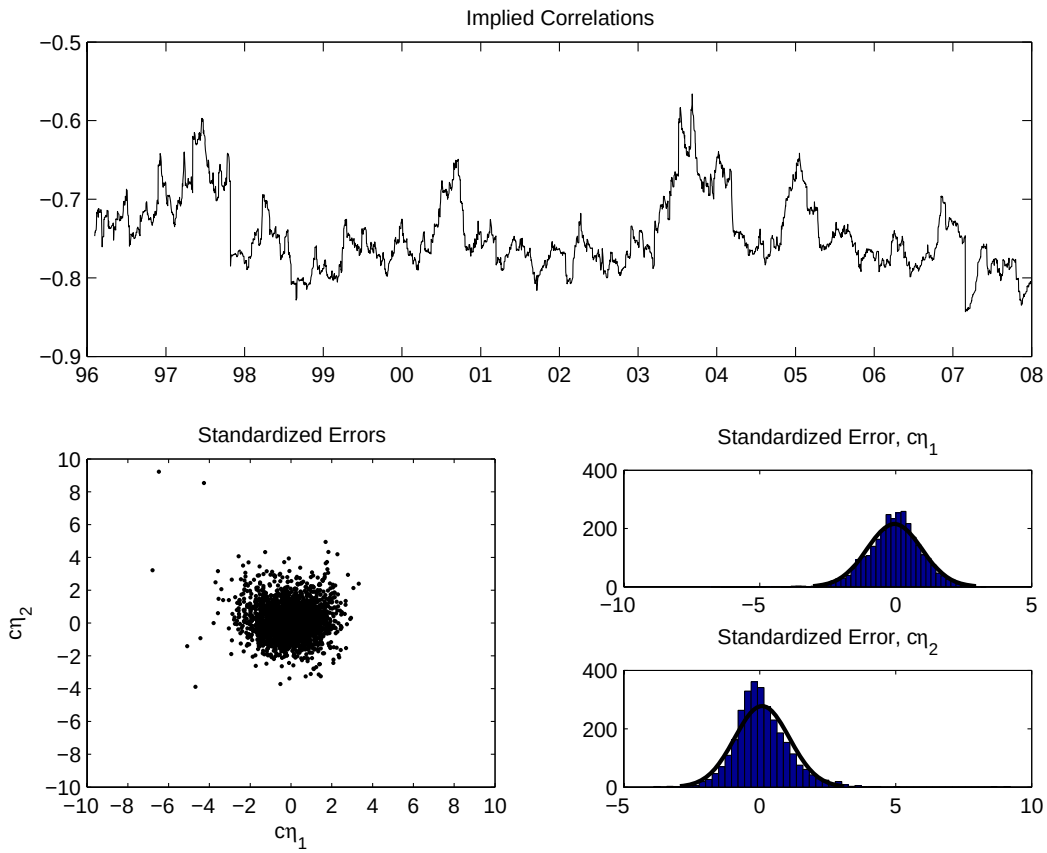


FIGURE 1.1: Estimated VAR-GARCH-DCC Model

The first panel plots the daily conditional correlations between the returns and the variance risk premium implied by the estimated VAR(1)-GARCH(1,1)-DCC model described in the main text. The lower left and right two panels provide a scatterplot and histograms, respectively, for the standardized residuals from the estimated model, $\widehat{c\eta}_t$. The daily sample used in estimating the model spans the period from February 1, 1996 to December 31, 2007, for a total of 2,954 daily observations.

Table 1.3: Country Specific Regressions

The results are based on the monthly regression in equation (2). t^{NW}-statistics are reported in parentheses. The sample period extends from January 2000 to December 2011.

Index	Horizon	1	2	3	4	5	6	9	12
AEX	Constant	-10.29	-10.55	-10.31	-10.73	-11.10	-10.79	-10.15	-9.74
		(-1.37)	(-1.44)	(-1.42)	(-1.49)	(-1.55)	(-1.49)	(-1.36)	(-1.29)
	VRP_t^i	0.16	0.17	0.13	0.16	0.18	0.15	0.08	0.05
		(0.98)	(1.37)	(1.51)	(1.91)	(3.04)	(2.86)	(1.94)	(1.47)
	$Adj.R^2$	0.14	0.91	0.58	1.69	3.20	2.25	0.50	-0.07
BEL 20	Constant	-4.11	-4.58	-4.85	-5.04	-5.20	-5.20	-4.93	-4.79
		(-0.67)	(-0.69)	(-0.68)	(-0.69)	(-0.70)	(-0.69)	(-0.63)	(-0.61)
	VRP_t^i	0.42	0.31	0.21	0.22	0.21	0.17	0.08	0.04
		(3.00)	(3.25)	(3.36)	(3.44)	(3.56)	(3.79)	(2.45)	(1.38)
	$Adj.R^2$	7.07	5.59	3.35	4.53	4.33	2.93	0.47	-0.44
CAC 40	Constant	-9.30	-9.38	-9.46	-9.77	-9.75	-9.40	-8.60	-8.20
		(-1.55)	(-1.56)	(-1.56)	(-1.58)	(-1.54)	(-1.46)	(-1.28)	(-1.19)
	VRP_t^i	0.24	0.23	0.22	0.23	0.19	0.14	0.06	0.04
		(2.22)	(2.68)	(3.49)	(4.45)	(4.75)	(3.33)	(1.54)	(1.03)
	$Adj.R^2$	1.50	2.78	4.15	5.55	4.17	2.32	-0.05	-0.33
DAX 30	Constant	-5.11	-5.73	-5.70	-6.18	-6.27	-5.41	-4.35	-4.14
		(-0.65)	(-0.75)	(-0.76)	(-0.84)	(-0.85)	(-0.72)	(-0.56)	(-0.53)
	VRP_t^i	0.02	0.17	0.18	0.25	0.25	0.16	0.08	0.11
		(0.13)	(0.89)	(1.38)	(2.11)	(2.74)	(1.79)	(1.64)	(2.41)
	$Adj.R^2$	-0.71	0.19	0.75	2.70	3.28	1.12	-0.08	0.83
FTSE 100	Constant	-5.23	-5.71	-6.29	-6.71	-6.69	-6.37	-5.60	-5.33
		(-1.08)	(-1.20)	(-1.38)	(-1.46)	(-1.42)	(-1.32)	(-1.12)	(-1.04)
	VRP_t^i	0.04	0.07	0.15	0.18	0.16	0.13	0.04	0.02
		(0.39)	(1.03)	(3.26)	(4.19)	(4.14)	(2.45)	(1.15)	(0.44)
	$Adj.R^2$	-0.65	-0.36	1.56	3.64	3.40	2.25	-0.30	-0.70
Nikkei 225	Constant	-7.45	-7.94	-8.55	-8.38	-8.07	-7.46	-6.16	-5.57
		(-1.11)	(-1.22)	(-1.31)	(-1.32)	(-1.27)	(-1.17)	(-0.93)	(-0.83)
	VRP_t^i	-0.01	0.02	0.11	0.12	0.11	0.09	0.02	0.02
		(-0.11)	(0.18)	(1.13)	(1.38)	(1.38)	(1.13)	(0.41)	(0.46)
	$Adj.R^2$	-0.71	-0.70	0.08	0.60	0.50	0.24	-0.65	-0.70
SMI	Constant	-2.73	-3.78	-4.01	-4.84	-5.27	-5.07	-4.45	-4.36
		(-0.46)	(-0.65)	(-0.69)	(-0.85)	(-0.93)	(-0.89)	(-0.76)	(-0.72)
	VRP_t^i	0.03	0.12	0.14	0.22	0.24	0.20	0.13	0.12
		(0.22)	(1.08)	(1.35)	(2.12)	(2.97)	(2.56)	(2.20)	(3.01)
	$Adj.R^2$	-0.69	-0.08	0.42	2.98	4.11	3.08	1.36	1.62
S&P 500	Constant	-6.64	-6.25	-6.34	-6.09	-6.17	-5.31	-4.12	-3.68
		(-1.46)	(-1.33)	(-1.36)	(-1.26)	(-1.24)	(-1.04)	(-0.77)	(-0.68)
	VRP_t^i	0.50	0.38	0.37	0.34	0.30	0.20	0.06	0.03
		(4.17)	(4.36)	(6.39)	(5.37)	(5.13)	(3.26)	(1.30)	(0.61)
	$Adj.R^2$	8.89	8.72	13.03	12.83	10.77	5.26	0.10	-0.53

Table 1.4: "Global" Variance Risk Premium Regressions

The results are based on the monthly regression in equation (3). t^{NW}-statistics are reported in parentheses. The sample period extends from January 2000 to December 2011.

Index	Horizon	1	2	3	4	5	6	9	12
AEX	Constant	-11.38	-11.52	-11.22	-11.52	-12.09	-11.39	-10.06	-9.55
		(-1.56)	(-1.59)	(-1.58)	(-1.61)	(-1.68)	(-1.56)	(-1.34)	(-1.26)
	VRP_t^{global}	0.29	0.28	0.23	0.25	0.29	0.21	0.06	0.02
		(1.99)	(2.54)	(3.45)	(3.98)	(4.72)	(2.92)	(1.11)	(0.48)
	$Adj.R^2$	0.80	1.80	1.77	2.86	4.49	2.51	-0.36	-0.70
BEL 20	Constant	-8.35	-7.78	-7.47	-7.58	-7.33	-6.72	-5.58	-5.05
		(-1.18)	(-1.09)	(-1.04)	(-1.05)	(-0.99)	(-0.89)	(-0.72)	(-0.65)
	VRP_t^{global}	0.42	0.33	0.29	0.31	0.26	0.19	0.08	0.03
		(2.79)	(3.50)	(4.70)	(6.87)	(5.26)	(3.06)	(1.53)	(0.63)
	$Adj.R^2$	3.77	3.62	3.81	5.67	3.83	1.95	-0.16	-0.66
CAC 40	Constant	-11.09	-11.18	-11.12	-11.29	-11.24	-10.38	-8.96	-8.31
		(-1.81)	(-1.84)	(-1.85)	(-1.85)	(-1.77)	(-1.61)	(-1.34)	(-1.21)
	VRP_t^{global}	0.33	0.33	0.32	0.32	0.29	0.20	0.08	0.03
		(2.89)	(3.38)	(5.51)	(6.99)	(5.22)	(3.01)	(1.67)	(0.87)
	$Adj.R^2$	1.90	3.95	5.51	7.34	6.30	3.12	0.01	-0.57
DAX 30	Constant	-7.12	-7.31	-7.26	-7.49	-7.52	-6.27	-4.49	-3.86
		(-0.96)	(-0.98)	(-1.01)	(-1.03)	(-1.00)	(-0.82)	(-0.58)	(-0.49)
	VRP_t^{global}	0.29	0.33	0.33	0.37	0.35	0.23	0.08	0.05
		(2.21)	(2.23)	(4.59)	(6.41)	(3.70)	(2.25)	(1.20)	(0.86)
	$Adj.R^2$	0.58	2.36	4.05	6.47	6.61	2.98	-0.13	-0.50
FTSE 100	Constant	-6.30	-6.50	-6.73	-6.87	-7.03	-6.57	-5.71	-5.37
		(-1.41)	(-1.44)	(-1.53)	(-1.52)	(-1.51)	(-1.37)	(-1.15)	(-1.05)
	VRP_t^{global}	0.18	0.18	0.21	0.21	0.21	0.16	0.06	0.02
		(1.74)	(1.82)	(3.45)	(4.63)	(3.48)	(2.23)	(1.30)	(0.47)
	$Adj.R^2$	0.63	1.56	3.97	5.33	5.80	3.57	0.04	-0.65
Nikkei 225	Constant	-8.76	-8.69	-9.00	-8.62	-8.36	-7.51	-5.97	-5.37
		(-1.25)	(-1.29)	(-1.41)	(-1.37)	(-1.31)	(-1.17)	(-0.91)	(-0.81)
	VRP_t^{global}	0.16	0.13	0.23	0.24	0.20	0.14	0.01	0.00
		(1.21)	(1.74)	(3.88)	(4.45)	(2.84)	(1.67)	(0.15)	(0.13)
	$Adj.R^2$	-0.18	-0.09	2.03	2.81	2.12	0.91	-0.74	-0.77
SMI	Constant	-3.80	-4.80	-5.14	-5.44	-5.69	-5.23	-4.01	-3.71
		(-0.67)	(-0.87)	(-0.94)	(-0.98)	(-1.01)	(-0.91)	(-0.68)	(-0.61)
	VRP_t^{global}	0.17	0.24	0.27	0.29	0.27	0.21	0.06	0.03
		(1.62)	(3.71)	(4.94)	(5.21)	(6.86)	(4.56)	(1.78)	(0.85)
	$Adj.R^2$	0.45	2.86	5.16	7.38	7.60	4.65	-0.07	-0.60
S&P 500	Constant	-6.38	-6.13	-6.34	-6.26	-6.20	-5.42	-4.28	-3.87
		(-1.39)	(-1.30)	(-1.37)	(-1.32)	(-1.27)	(-1.08)	(-0.80)	(-0.71)
	VRP_t^{global}	0.47	0.36	0.38	0.35	0.31	0.22	0.09	0.05
		(3.84)	(4.22)	(5.89)	(5.94)	(4.60)	(2.94)	(1.54)	(1.04)
	$Adj.R^2$	6.32	6.74	11.10	12.05	9.94	5.47	0.57	-0.06

Table 1.5: Panel Regressions

The results are based on the monthly "global" and country-specific panel regressions in the equation (4) and (5), respectively. NW-based t-statistics are reported in the parentheses. The sample period extends from January 2000 to December 2011.

	"Global" Regressors							
Horizon	1	2	3	4	5	6	9	12
Constant	-7.90	-7.99	-8.04	-8.13	-8.18	-7.44	-6.13	-5.64
	(-3.08)	(-3.98)	(-4.56)	(-5.06)	(-5.29)	(-4.88)	(-4.63)	(-5.89)
VRP_t^{global}	0.29	0.27	0.28	0.29	0.27	0.19	0.06	0.03
	(5.23)	(7.27)	(7.58)	(10.91)	(8.44)	(6.33)	(3.26)	(1.72)
$Adj.R^2$	1.92	3.03	4.67	6.21	5.91	3.43	0.44	0.06
Constant	3.09	-1.20	0.17	7.60	9.71	9.67	6.96	-1.59
	(0.14)	(-0.06)	(0.01)	(0.56)	(0.70)	(0.72)	(0.72)	(-0.22)
VRP_t^{global}	0.30	0.28	0.29	0.31	0.28	0.21	0.07	0.03
	(5.95)	(9.41)	(9.97)	(13.58)	(10.79)	(9.04)	(5.19)	(2.30)
$\log(P_t/E_t)^{global}$	-5.07	-3.13	-3.78	-7.24	-8.23	-7.86	-5.99	-1.85
	(-0.49)	(-0.34)	(-0.52)	(-1.14)	(-1.28)	(-1.25)	(-1.32)	(-0.58)
$Adj.R^2$	1.88	2.97	4.65	6.38	6.21	3.74	0.66	0.00
	Country-Specific Regressors							
Horizon	1	2	3	4	5	6	9	12
Constant	-6.87	-7.03	-6.98	-7.24	-7.32	-6.89	-6.08	-5.73
	(-2.56)	(-2.72)	(-3.22)	(-3.52)	(-4.09)	(-4.29)	(-5.11)	(-5.27)
VRP_t^i	0.20	0.19	0.18	0.21	0.20	0.15	0.07	0.05
	(4.64)	(4.08)	(5.48)	(6.92)	(5.69)	(4.14)	(2.76)	(2.47)
$Adj.R^2$	1.17	1.88	2.57	4.06	4.25	2.77	0.79	0.50
Constant	2.31	3.15	2.80	2.97	3.15	2.93	2.48	0.33
	(0.40)	(0.57)	(0.54)	(0.68)	(0.81)	(0.78)	(0.87)	(0.15)
VRP_t^i	0.21	0.20	0.19	0.22	0.21	0.16	0.08	0.06
	(5.58)	(5.62)	(9.51)	(10.47)	(7.24)	(5.39)	(3.78)	(3.61)
$\log(P_t^i/E_t^i)$	-4.61	-5.11	-4.91	-5.12	-5.24	-4.91	-4.27	-3.02
	(-1.79)	(-2.22)	(-2.24)	(-2.63)	(-2.98)	(-2.82)	(-3.05)	(-2.62)
$Adj.R^2$	1.18	2.00	2.75	4.33	4.60	3.12	1.15	0.71

Table 1.6: Panel Regressions with Forecasted "Global" Variance Risk Premium

The results are based on the monthly forecasted "global" panel regressions in the equation (4) and (5), respectively. NW-based t-statistics are reported in the parentheses. The sample period expends from January 2000 to December 2011.

Horizon	1	2	3	4	5	6	9	12
Constant	-10.43	-9.24	-8.51	-8.54	-9.01	-8.70	-7.20	-6.39
	(-3.50)	(-3.58)	(-3.56)	(-3.90)	(-4.42)	(-4.61)	(-4.72)	(-5.37)
$FVRP_t^{global}$	0.49	0.35	0.28	0.28	0.32	0.30	0.17	0.10
	(6.05)	(6.26)	(5.19)	(5.98)	(6.19)	(5.07)	(3.75)	(2.49)
$Adj.R^2$	4.86	4.26	3.91	4.80	7.68	7.63	3.22	1.56
Constant	4.61	-2.52	-3.14	3.16	10.34	12.38	10.32	0.89
	(0.19)	(-0.12)	(-0.18)	(0.21)	(0.73)	(1.02)	(1.03)	(0.13)
$FVRP_t^{global}$	0.50	0.35	0.28	0.29	0.34	0.32	0.18	0.11
	(6.79)	(7.47)	(7.01)	(8.02)	(7.87)	(6.24)	(5.46)	(3.42)
$\log(P_t/E_t)^{global}$	-6.95	-3.10	-2.48	-5.39	-8.90	-9.69	-8.03	-3.33
	(-0.62)	(-0.32)	(-0.30)	(-0.76)	(-1.35)	(-1.67)	(-1.77)	(-1.14)
$Adj.R^2$	4.86	4.21	3.85	4.86	8.05	8.15	3.70	1.59

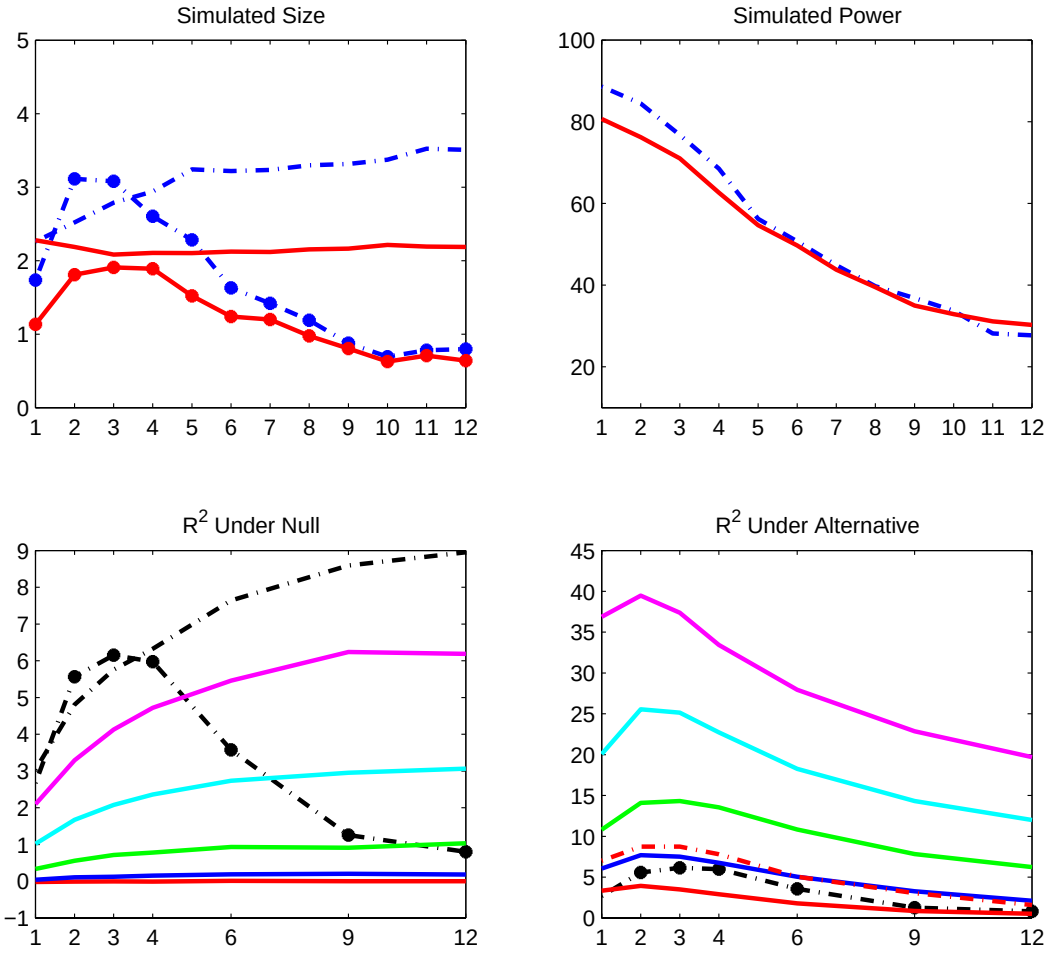


FIGURE 1.2: Simulated Size and Power

The upper left panel reports the 95-percentiles in the finite-sample distributions of the t^{NW} (dash line) and t^{HD} (solid line) based on simulated "monthly" data from the restricted VAR-GARCH-DCC model under the null of no predictability. The dashed and solid star lines refer to the corresponding t-statistics for actual monthly U.S. S&P 500 returns spanning Feb1996 to Dec2007. The lower left panel plots the quantiles in the finite-sample distribution of the R^2 from the return regression in equation (1) and simulated "monthly" date from the restricted VAR-GARCH-DCC model under the null of no predictability. The star dashed line refer to the corresponding R^2's in actual daily U.S. S&P 500 returns spanning February 1, 1996 to December 31, 2007. The right two panels are based on unrestricted VAR-GARCH-DCC model: the top panel give simulated "monthly" percentage power and the size-adjusted 5-percent t^{NW} (dashed line) and t^{HD} (solid line) statistics; the bottom panel reports the quantiles in the simulated finite-sample distribution.

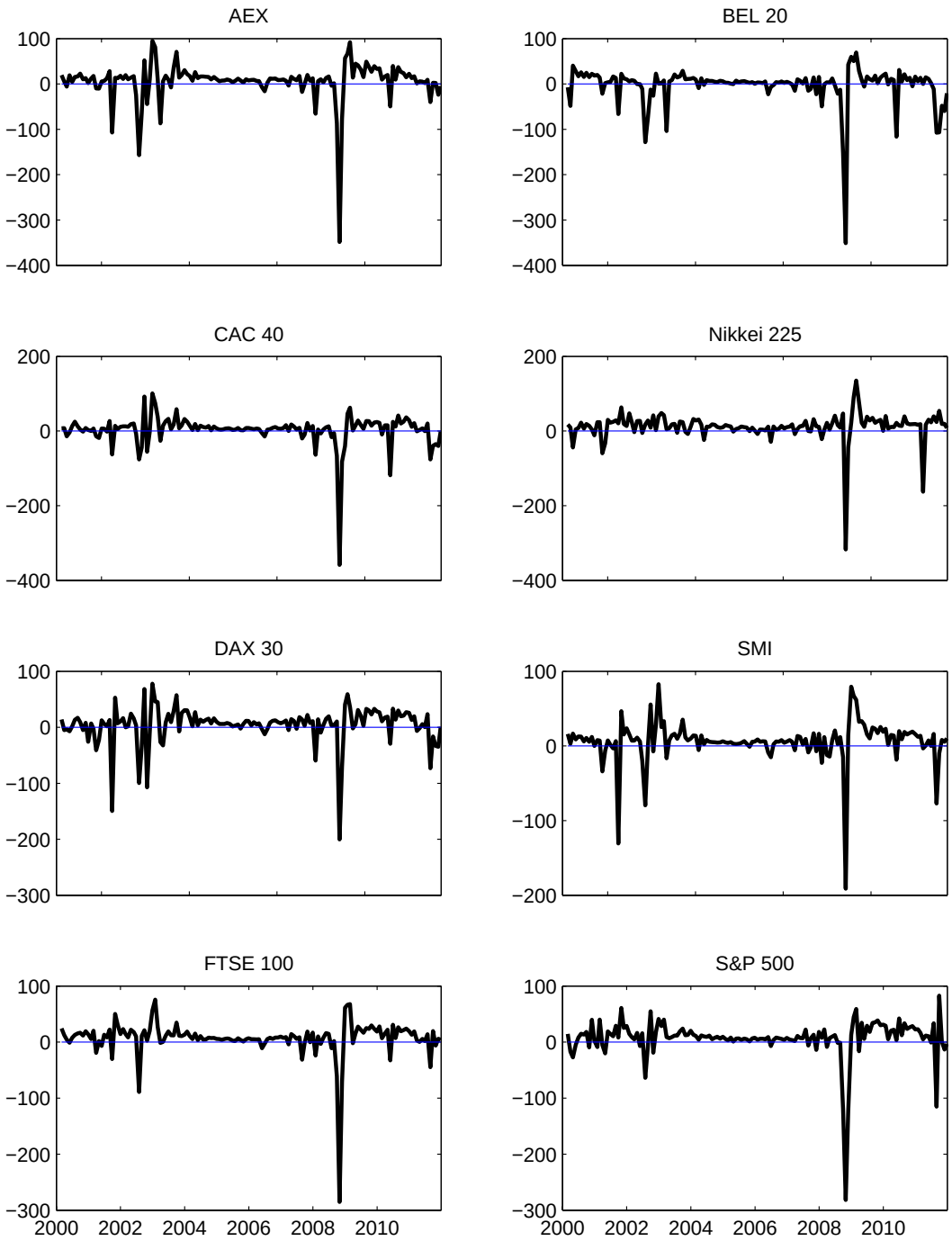


FIGURE 1.3: Variance Risk Premia

The figure shows the monthly proxies for the variance risk premia VRP_t^i for Netherlands (AEX), Belgium(BEL 20), France (CAC 40), Japan (Nikkei 225), Germany (DAX 30), Switzerland (SMI 20), the U.K. (FTSE 100), and the U.S. (S&P 500). The risk premia are constructed by subtracting the actual realized variation from the model-free options implied variation. The sample period spans January 2000 to December 2011

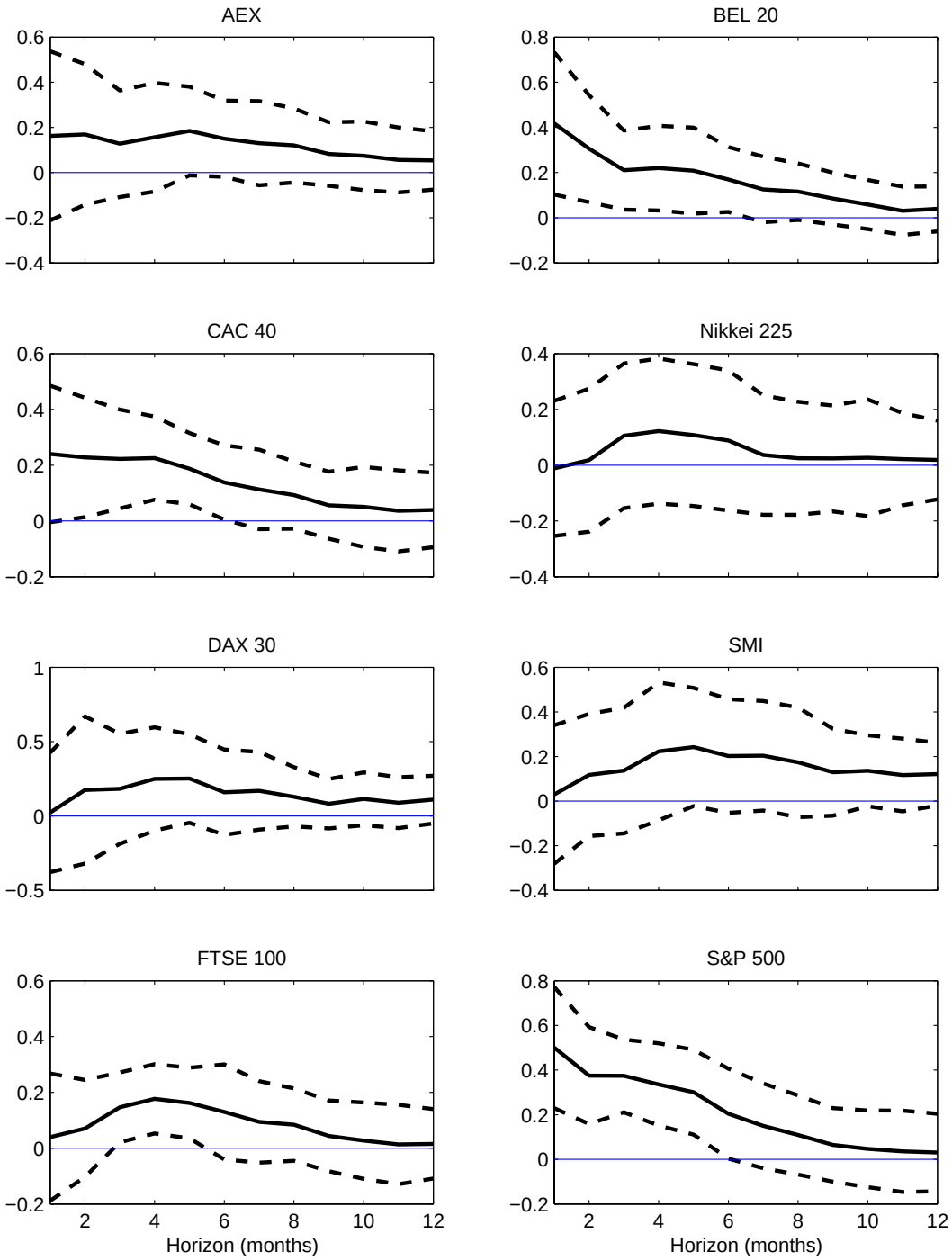


FIGURE 1.4: Country Specific Regression Coefficients

The figure shows the estimated regression coefficients for VRP_t^i for each of the country specific return regressions reported in Table 2.4, together with NW-based 90% standard error bands, see Table 1 for simluated critical value from one to twelve months. The regressions are based on monthly data from January 2000 to December 2011.

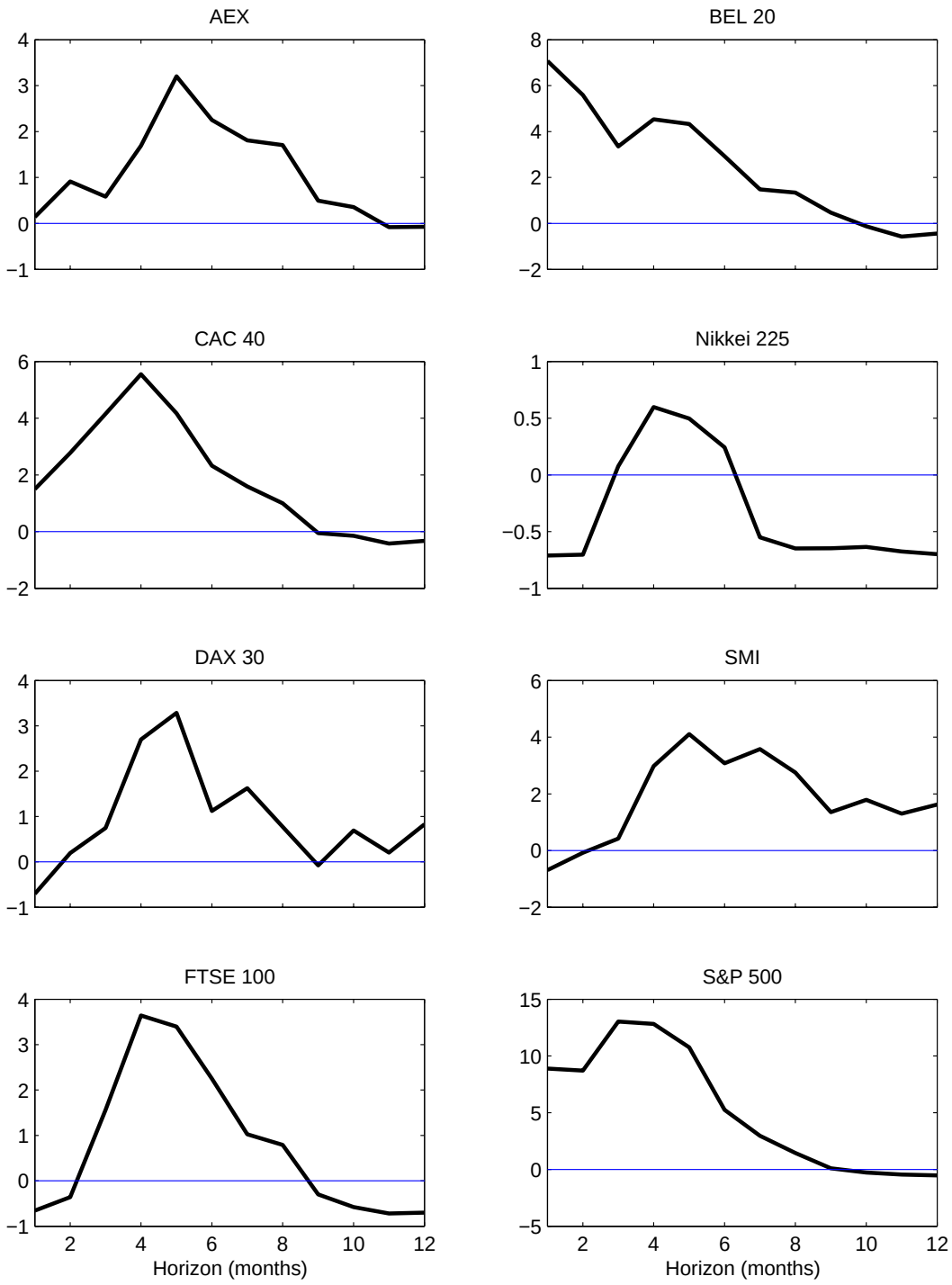


FIGURE 1.5: Country Specific Regression R^2's

The figure shows the adjusted $R^2(h)$'s for the country specific return regressions reported in Table 2.4. The regressions are based on monthly data from January 2000 to December 2011.

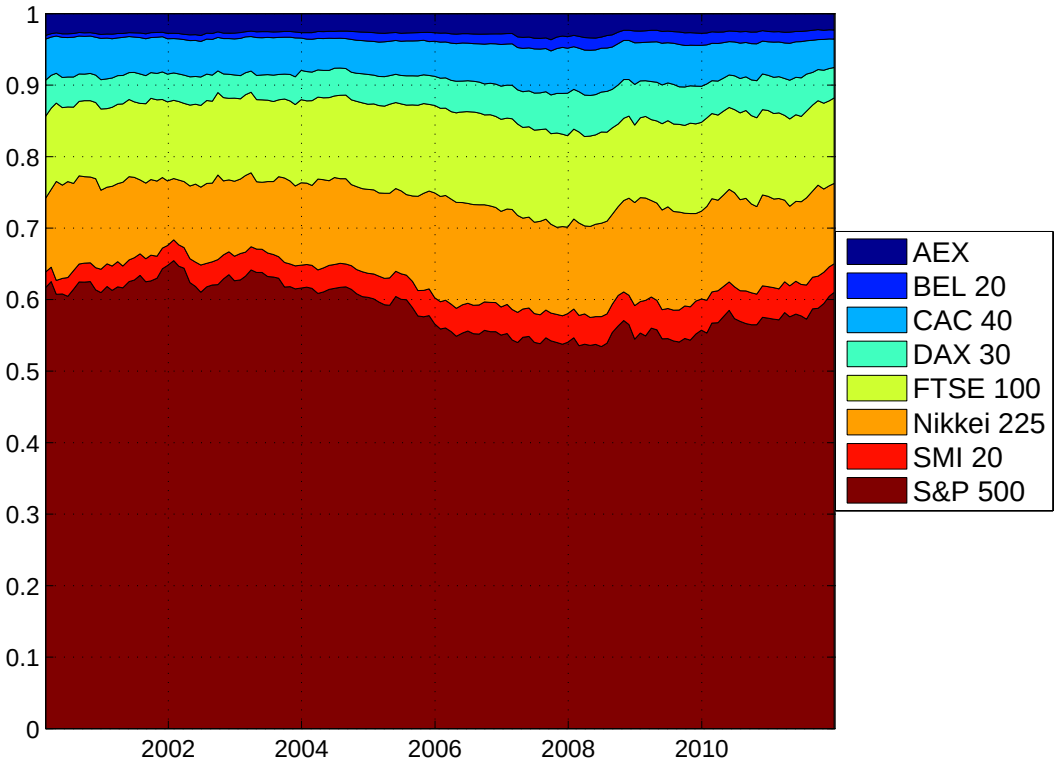


FIGURE 1.6: Market Capitalization

The figure shows the relative market capitalization by aggregate index for Netherlands (AEX), Belgium(BEL 20), France (CAC 40), Germany (DAX 30), the U.K. (FTSE 100), Japan (Nikkei 225), Switzerland (SMI 20), and the U.S. (S&P 500).

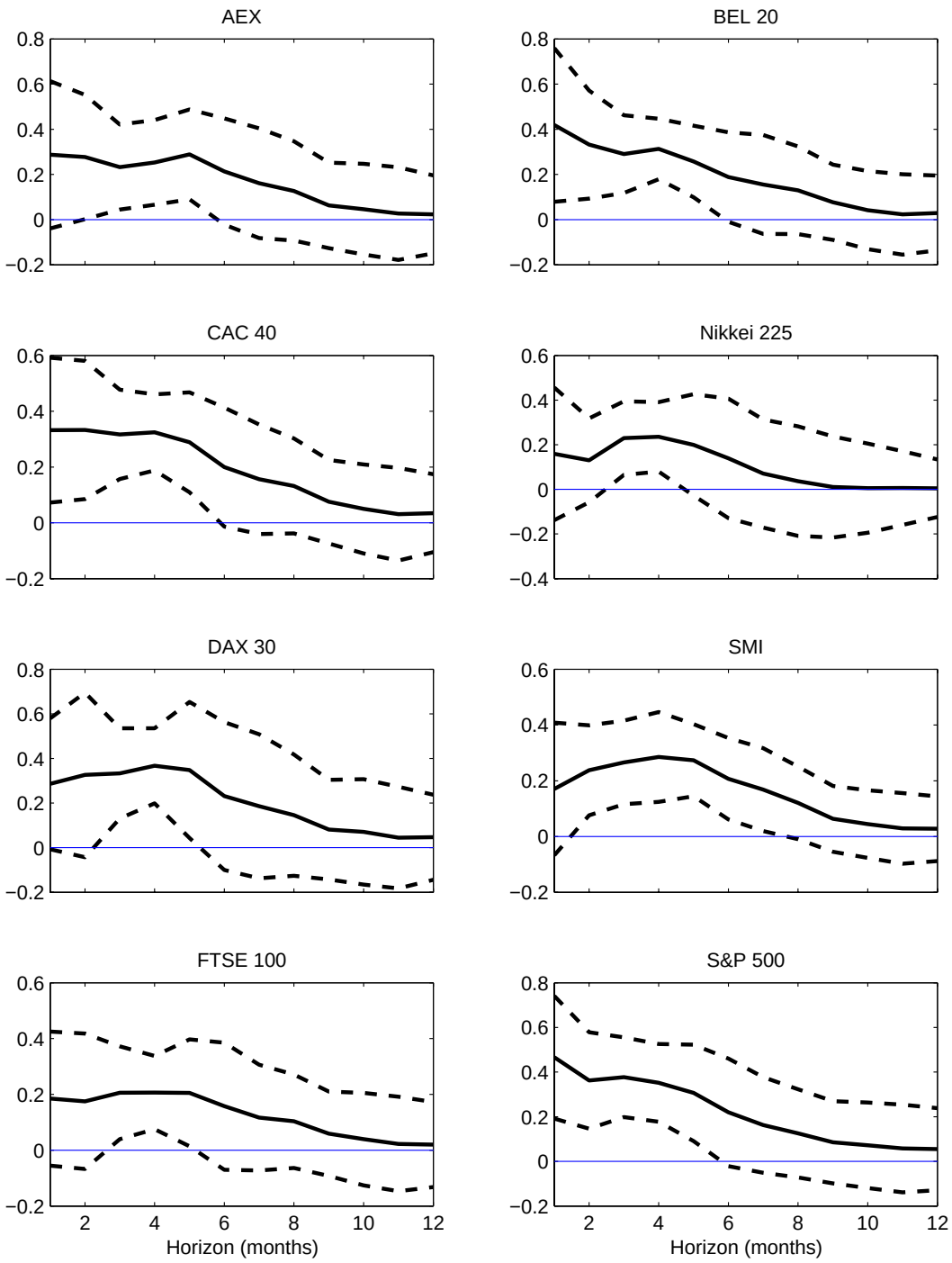


FIGURE 1.7: "Global" VRP Regression Coefficients

The figure shows the coefficient estimates for VRP_t^{global} from the return regressions reported in Table A, together with NW-based 90% standard error bands, see Table 1 for simluated critical value from one to twelve months. The regressions are based on monthly data from January 2000 to December 2011.

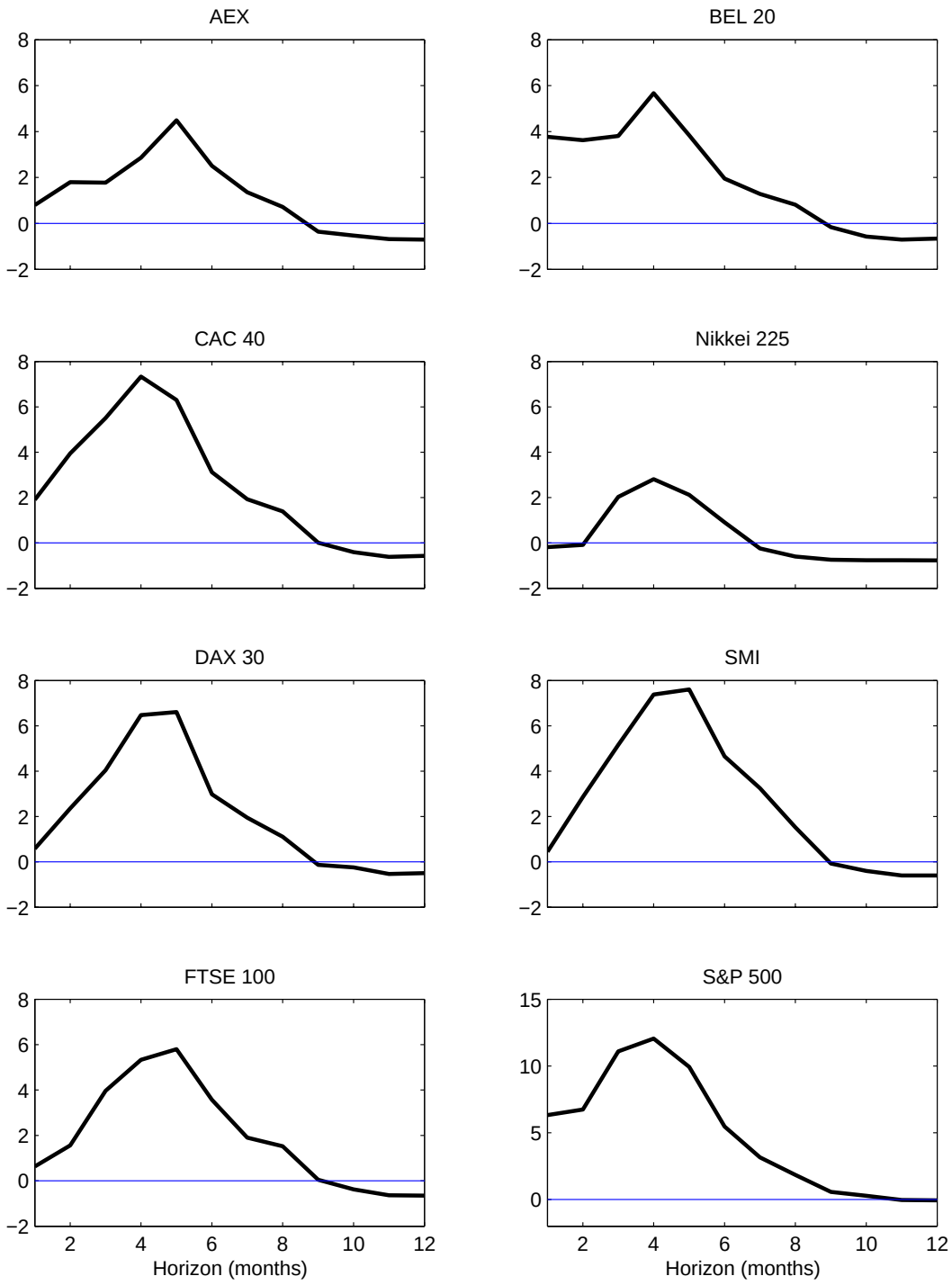


FIGURE 1.8: "Global" VRP Regression R^2's

The figure shows the adjusted $R^2(h)$'s from regressing the individual country returns on VRP_t^{global} reported in Table A. The regressions are based on monthly data from January 2000 to December 2011.

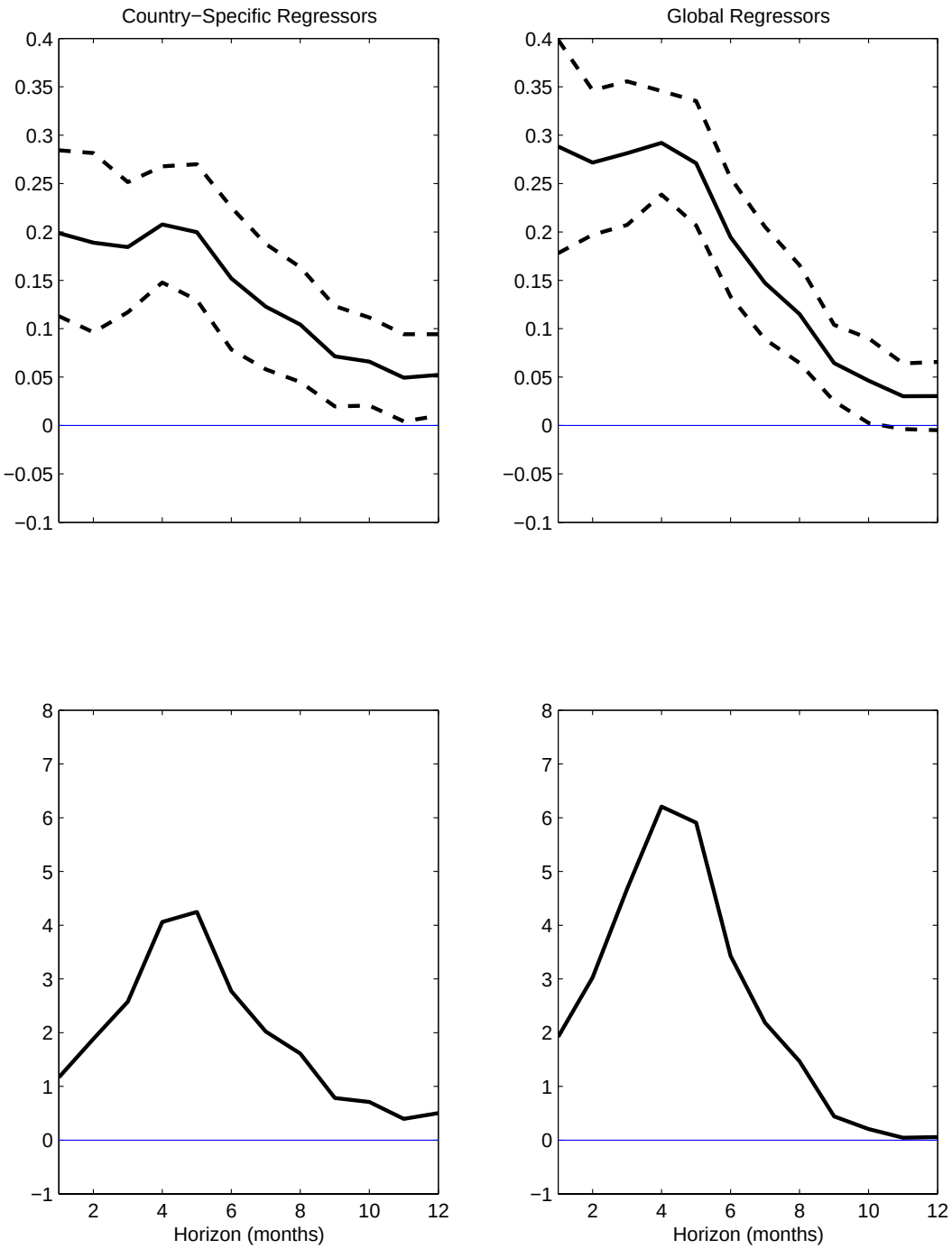


FIGURE 1.9: Panel Regression Coefficients and R^2's

The top two panels show the estimated panel regression coefficients from regressing the returns on the individual country variance risk premia VRP_t^i and the "global" variance risk premium VRP_t^{global}, respectively, reported in Table A, together with two NW-based standard error bands. The bottom two panels show the $R^2(h)$'s from the same two panel regressions. The regressions are based on monthly data from January 2000 through December 2011.

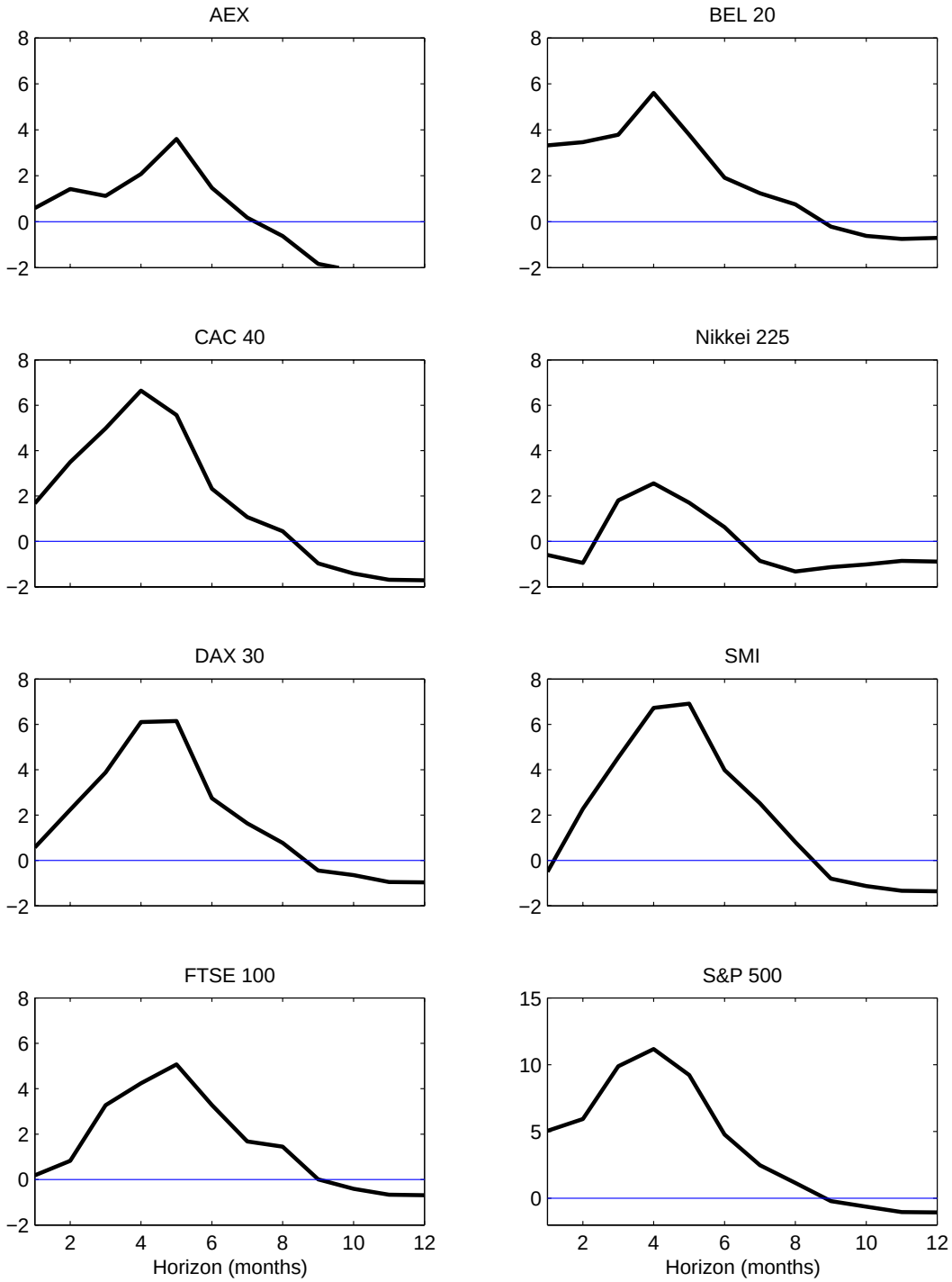


FIGURE 1.10: "Global" VRP Panel Regression R^2's

The figure shows the adjusted $R^2(h)$'s implied by the VRP_t^{global} panel regressions reported in the top panel in Table A. The regressions are based on monthly data from January 2000 to December 2011.

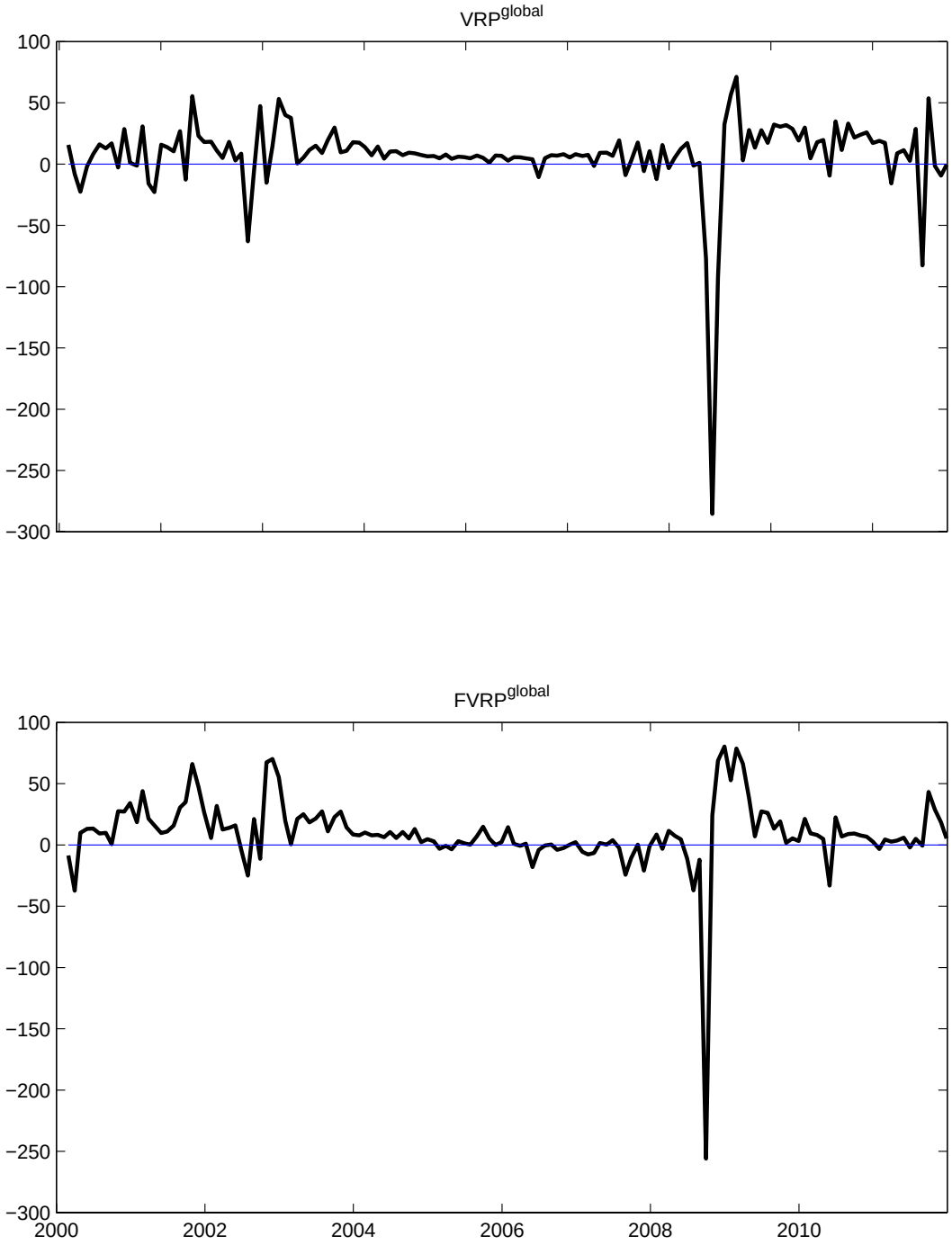


FIGURE 1.11: Variance Risk Premia

The figure shows our proxies for the monthly "global" variance risk premia VRP_t^{global} (top panel) and $FVRP_t^i$ (bottom panel) as defined in the main text. The sample period spans January 2000 to December 2011

2

Stock Return and Cash Flow Predictability: The Role of Volatility Risk

2.1 Introduction

Counter to the "old" efficient market hypothesis dictum that speculative returns are largely unpredictable over time, it is now generally accepted that equity returns are both time-varying and predictable. It is also widely believed that the predictability of the aggregate stock market as a whole is the strongest over longer multi-year horizons.[1] At the same time, to the extend that a consensus has emerged it suggests that expected dividend growth rates for the aggregate market portfolio, or aggregate cash flows, are much less predictable than the expected returns.[2]

[1] Some of the predictor variables used in establishing long-run return predictability include: dividend-price, earning-price, and other valuation ratios (Campbell and Shiller, 1988a,c; Fama and French, 1988b; Lamont, 1998; Lewellen, 2004); firms' net equity payout (Boudoukh, Richardson, and Roberts, 2007) and equity issuance (Baker and Wurgler, 2000); interest-rate variables such as t-bill and t-bond rates, term spreads, and default spreads (Campbell, 1987; Fama and French, 1989; Hodrick, 1992); and macroeconomic variables like total investment (Cochrane, 1991), the consumption-wealth ratio (Lettau and Ludvigson, 2001), and inflation (Campbell and Vuolteenaho, 2004).

[2] Maio and Santa-Clara (2013) have recently challenged this view, showing that for portfolios comprised of small and value stocks, the dividend-price ratio is primarily related to future changes in cash flows. With a few notable exceptions (e.g., Fama and French, 1988b; Lettau and Ludvigson,

Much of the literature underlying these findings, and the choice of predictor variables in particular, have been guided by the present-value framework pioneered by Campbell and Shiller (1988a,c), and the implication that the dividend-price ratio, or the dividend yield, is identically equal to the expected value of the future returns discounted by the future dividend growth rates. As emphasized by Cochrane (2008, 2011), this intimate link between dividend growth and stock return predictability also implies that the seemingly stronger empirical evidence for long-run return predictability is not surprisingly accompanied by seemingly weaker empirical evidence for long-run dividend growth predictability.

Set against this background, a number of recent studies have argued that the variance risk premium, or the difference between options implied and expected variances, possesses superior forecasting power for stock market returns over shorter within-year horizons; see, e.g., Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011b), and Kelly (2011). Motivated by these more recent empirical findings, we show how explicitly incorporating priced volatility risk into the present-value framework affords important new insights into the return vis-à-vis dividend growth predictability debate across *all* horizons.

The reduced form VAR framework, as exemplified by Hodrick (1992) and Campbell (2001), traditionally used for empirically implementing present value relations does not naturally lend itself to the estimation of models involving priced volatility risk. Instead, we follow Sentana and Fiorentini (2001) and Rigobon (2003) in designing a "structural" factor GARCH model, in which the factors exhibit time-varying volatility. The dynamics of the factors is derived endogenously from an extended long-run risk model explicitly incorporating time-varying consumption volatility and volatility-of-volatility, or economic uncertainty. The resulting econometric model separately identifies the long-run risk, volatility, and economic uncertainty compo-

2005) cash flow predictability has historically received much less attention in the literature.

nents, as well as the corresponding structural shocks and their contemporaneous impact on both returns and dividend growth.

Estimating the "structural" factor GARCH model by standard GMM techniques on data for the S&P 500 market portfolio, we confirm existing empirical evidence that the dividend-price ratio is useful for predicting long-horizon multi-year returns, but that it has no predictive power for dividend growth.[3] More important, we document a number of new results pertaining to the predictability of the volatility factors. In particular, while the variance risk premium shows significant predictability for returns over short within-year horizons, it also helps predict dividend growth. Similarly, the expected return variation appears to be very informative for predicting dividend growth.

These results are consistent with the findings in Koijen and van Nieuwerburgh (2011) that the high-frequency component of the dividend-price ratio, which in our setup is driven by two separate volatility factors, contains useful information for predicting expected dividend growth. Our results are also related to Binsbergen, Brandt, and Koijen (2012) and their findings that the term structure of equity risk premia is particularly steep in the short end, while standard asset pricing models without priced volatility risk typically imply higher equity premia at the long end.

In addition to the new empirical evidence pertaining to the short-run predictability of returns and dividend growth, by explicitly identifying the systematic risk factors at work, our "structural" factor GARCH approach also helps shed new light on the underlying economic mechanisms. Specifically, we find that the long-run expected growth component is highly persistent with a first-order autocorrelation coefficient close to one ($\rho_x = 0.988$) at the monthly level, consistent with the idea

[3] Compared to earlier empirical findings based on univariate regressions (Rozeff, 1984; Fama and French, 1988b; Campbell and Shiller, 1988c) and traditional present-value homoskedastic VAR's (Hodrick, 1992; Campbell, 2001; Cochrane, 2008), our "structural" factor GARCH model results in much sharper inference, with the actual point estimates systematically falling within the standard error bands obtained from the more conventional procedures.

in Bansal and Yaron (2004a) that it acts as the most important driver of the risk premium dynamics over long horizons.[4] The model also clearly differentiates and is able to accurately estimate the persistence of the consumption volatility component ($\rho_\sigma = 0.64$) and the volatility-of-volatility, or economic uncertainty, component ($\rho_q = 0.46$), advocated by Bollerslev, Tauchen, and Zhou (2009), both of which are intimately linked to the shorter-run predictability patterns in the data. In terms of the underlying "structural" shocks, we find a negative relationship between the long-run growth and consumption volatility shocks (akin to a "leverage effect"), as well as a negative relationship between the consumption volatility and volatility uncertainty shocks (interpretable as a separate new "leverage effect"). The price-dividend ratio also responds negatively to both consumption volatility and volatility uncertainty shocks.[5]

The basic motivation behind the new "structural" factor GARCH model is in line with a growing recent literature seeking to explicitly incorporate the effect of stochastic volatility in asset pricing models. For example, Bansal, Kiku, Shaliastovich, and Yaron (2013) demonstrate that ignoring the variation in volatility leads to counter-intuitive economic interpretation of risk premium dynamics. Similarly, Campbell, Giglio, Polk, and Turley (2013) examine the cross-sectional return predictability in an ICAPM framework that allows for stochastic volatility.[6] In contrast to these studies, our focus is on the *joint* predictability of returns and cash flows within the context of a "structural" econometric model explicitly designed to accom-

[4] Nakamura, Sergeyev, and Steinsson (2012) have recently shown how the long-run growth factor may also be identified from cross-country aggregate consumption data under additional simplifying assumptions.

[5] The importance of economic uncertainty for explaining asset prices has also recently been emphasized from different perspectives by Bekaert, Engstrom, and Xing (2009), Nieto and Rubio (2011), and Corradi, Distaso, and Mele (2013), among others.

[6] Our "structural" factor GARCH estimate for the persistence in consumption volatility ρ_σ, and in turn the effect of allowing for time-varying volatility, are much larger than the estimates reported in Campbell, Giglio, Polk, and Turley (2013) based on simple VAR procedures and imprecise variance measures.

modate time-varying volatility in an internally consistent fashion. Recent studies by Binsbergen and Koijen (2010) and Piatti and Trojani (2012) have also relied on a latent variable approach with heteroskedastic shocks for incorporating the effect of time-varying volatility within a present-value framework. Importantly, however, we differ from both of these studies by specifying an empirically more realistic two-factor volatility structure and by explicitly including both the actual and risk-neutral expected variation in the formulation and estimation of the model.[7]

The rest of the paper is organized as follows. Section 2 presents the equilibrium asset pricing model underlying our empirical investigations. Section 3 describes the data and the formulation of the "structural" factor GARCH model and the GMM-based parameter estimation results. Section 4 details the return and cash flow predictability implied by the model, and contrast the results with those obtained by other less structured reduced form estimation procedures. Section 5 concludes.

2.2 Asset Pricing Model

Our equilibrium-based approach combines the long-run risk model pioneered by Bansal and Yaron (2004a), with the model in Bollerslev, Tauchen, and Zhou (2009) explicitly allowing for stochastic volatility-of-volatility, or time-varying economic uncertainty. This general setup naturally accommodates the magnitude of both the equity and variance risk premia, as well as the long- and short-horizon predictability patterns in the returns and cash flows within a unified framework.

2.2.1 Model Setup and Assumptions

Following the long-run risk literature, we assume an endowment economy with a representative agent equipped with Epstein and Zin (1991) recursive preferences. The

[7] Other recent studies seeking to incorporate more realistic two-factor volatility structures in the standard long-run risk model include Zhou and Zhu (2013), Branger and Vòlkert (2012), and Branger, Rodriguez, and Schlag (2011), among others.

logarithm of the intertemporal marginal substitution for this agent may consequently be expressed as,

$$m_{t+1} = \theta \log \delta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1) r_{c,t+1}, \tag{2.1}$$

where $r_{c,t+1} \equiv \log(R_{c,t+1})$ refers to the logarithmic return on the consumption asset, $\Delta c_{t+1} \equiv \log(C_{t+1}/C_t)$ denotes the growth rate of consumption, $0 < \delta < 1$ is the time discount factor, $\gamma > 0$ denotes the risk aversion parameter, and $\theta \equiv \frac{1-\gamma}{1-\psi^{-1}}$ where $\psi > 0$ refers to the intertemporal elasticity of the substitution. As is standard in the long-run risk literature, we will assume that $\gamma > 1$, implying that the representative agent is more risk averse than log utility, and that $\psi > 1$, and therefore $\theta < 0$, implying a preference for early resolution of uncertainty.

Let x_t denote the long-run mean of consumption growth as in Bansal and Yaron (2004a), and σ_t^2 and q_t refer to two separate volatility factors along the lines of Bollerslev, Tauchen, and Zhou (2009). For notational convenience, collect the consumption growth Δc_t, the log dividend growth Δd_t, and the latent state variables in the vector $Y_t = (\Delta c_t, x_t, \sigma_t^2, q_t, \Delta d_t)'$. The importance of allowing for multiple volatility factors in accurately describing both short- and long-horizon time-varying return and volatility dynamics has also recently been highlighted by Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011b), Bollerslev, Sizova, and Tauchen (2012a), Zhou and Zhu (2013), Branger and Vòlkert (2012), among others.

We will assume that the state vector Y_t has affine conditional mean and variance dynamics,

$$Y_{t+1} = \mu + F Y_t + H G_t z_{t+1}, \tag{2.2}$$

where $z_{t+1} \equiv [z_{c,t+1}, z_{x,t+1}, z_{\sigma,t+1}, z_{q,t+1}, z_{d,t+1}]'$ denotes a vector of independent standard normally distributed shocks. We rank all of the "structural" consumption shocks, including the two volatility shocks $z_{\sigma,t}$ and $z_{q,t}$, before shocks to dividends

$z_{d,t}$. Based on the intuition that level shocks are more "fundamental" than shocks to volatility, we also put the $z_{c,t}$ and $z_{x,t}$ shocks before the two volatility shocks. The conditional mean of Y_t is in turn determined by the constant vector μ and the loading matrix F. We assume that this loading matrix takes the sparse form,

$$F = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & \rho_x & 0 & 0 & 0 \\ 0 & 0 & \rho_\sigma & 0 & 0 \\ 0 & 0 & 0 & \rho_q & 0 \\ 0 & \phi_{dx} & 0 & 0 & \rho_d \end{pmatrix}, \tag{2.3}$$

in which the diagonal elements characterize the own lagged dependencies and the off-diagonal elements describe the dynamic first-order cross dependencies. In particular, ϕ_{dx} allows the dividend growth rate Δd_{t+1} to directly load on the lagged long-run consumption growth component x_t. Allowing Δd_{t+1} to also depend on its own lag permits a non-redundant pricing effect of dividend growth risk on the equity premium. Restricting this coefficient ρ_d to be zero reduces the model's growth dynamics to that of a "standard" long-run risk model. However, our estimates of the model discussed below strongly rejects such a specification.

The conditional second-order dynamics of the state vector is determine by the time-varying diagonal volatility matrix G_t and the constant loading matrix H,

$$G_t = \begin{pmatrix} \sigma_t & 0 & 0 & 0 & 0 \\ 0 & \sqrt{q_t} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{q_t} & 0 & 0 \\ 0 & 0 & 0 & \sqrt{q_t} & 0 \\ 0 & 0 & 0 & 0 & \sigma_t \end{pmatrix} \qquad H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \varphi_x & 0 & 0 & 0 \\ 0 & \varphi_x s_{\sigma,x} & 1 & 0 & 0 \\ 0 & \varphi_x s_{q,x} & s_{q,\sigma} & \varphi_q & 0 \\ 0 & \varphi_x s_{d,x} & s_{d,\sigma} & \varphi_q s_{d,q} & \varphi_d \end{pmatrix}. \tag{2.4}$$

Our choice of G_t differs from the models in Drechsler and Yaron (2011b) and Branger and Vòlkert (2012) by allowing both x_{t+1} and σ^2_{t+1} to have time-varying volatility $\sqrt{q_t}$. Our choice of G_t also nests the model in Bollerslev, Tauchen, and Zhou (2009) by zeroing out the long-run growth component, equating the dividend and consumption

growth, and fixing $s_{i,j} = 0$ for $i \neq j$, thereby rendering H diagonal.[8] Identification of the lower triangular volatility loading matrix H is effectively accomplished through heteroskedasticity, and cross-dependencies between the different state variables implied by the form of the time-varying volatility.

Further, denoting the columns of $H \equiv [h_1, h_2, h_3, h_4, h_5]$, the "square" of HG_t may be conveniently expressed in affine form as,

$$HG_tG_t'H' = \sum_{j=1,5} h_j h_j' \sigma_t^2 + \sum_{j=2,3,4} h_j h_j' q_t. \tag{2.5}$$

This two-factor volatility structure is distinctly different from the one-factor setup recently employed in Campbell, Giglio, Polk, and Turley (2013). As discussed in more detail below, it affords an empirically much more realistic description of the return and cash flow dynamics, and in turn the predictability patterns obtained by imposing the equilibrium-based restrictions.

2.2.2 Model Implications

In order to deduce the "structural" model restrictions that guide our empirical analysis, we begin by solving the consumption-based asset pricing model using similar techniques to the ones in Bansal and Yaron (2004a), Bansal, Kiku, and Yaron (2007a), and Drechsler and Yaron (2011b). In the spirit of Campbell (1993, 1996), we then substitute out the hard-to-measure consumption and its volatility dynamics with directly observable market return and its variance measures.

Standard solution methods applied in the long-run risk literature readily imply

[8] We also experimented with two alternative setups, one closer to Drechsler and Yaron (2011b) with $G_t = \text{diag}[\sigma_t, \sqrt{q_t}, \sigma_t, \sqrt{q_t}, \sigma_t]$, and the other one closer to Branger and Vòlkert (2012) with $G_t = \text{diag}[\sigma_t, \sigma_t, \sigma_t, \sqrt{q_t}, \sigma_t]$, resulting in qualitatively similar predictability results to the ones reported below. However, both of these alternative specifications were rejected at conventional significance levels by the corresponding GMM-based J-tests for over-identifying restrictions. Further details concerning these alternative models and empirical results are reported in the supplementary Appendix.

that the stochastic discount factor m_{t+1}, the return on consumption $r_{c,t+1}$, and the market return on dividends $r_{t,t+1}$, must satisfy

$$\begin{array}{rcl} m_{t+1} - E_t(m_{t+1}) & = & -\Lambda' H G_t z_{t+1}, \\ r_{c,t+1} - E_t(r_{c,t+1}) & = & \Lambda_c' H G_t z_{t+1}, \\ r_{t,t+1} - E_t(r_{t,t+1}) & = & \Lambda_d' H G_t z_{t+1}, \end{array} \tag{2.6}$$

where $\Lambda = \gamma e_1 + \kappa_1(1-\theta)A$, for $A = (0, A_x, A_\sigma, A_q, 0)$, denotes the price of risk for the factor shocks, $\Lambda_c = e_1 + \kappa_1 A$, $\Lambda_d = e_5 + \kappa_{d,1} A_d$, κ_1 and $\kappa_{d,1}$ refer to the Campbell and Shiller (1988c) log-linearization constants based on the "usual" approximations for consumption return $r_{c,t+1} \approx \kappa_0 + \kappa_1 \nu_{t+1} - \nu_t + \Delta c_{t+1}$ and the aggregate market return $r_{t,t+1} \approx \kappa_{d,0} + \kappa_{d,1} w_{t+1} - w_t + \Delta d_{t+1}$, respectively, and the two selection vectors are defined by $e_1 \equiv [1,0,0,0,0]'$ and $e_5 \equiv [0,0,0,0,1]'$.[9] Given these expressions, it is possible to solve for the market return variance $Var_t(r_{t,t+1})$, the variance risk premium VRP_t, and the log dividend-price ratio dp_t, as

$$Var_t(r_{t,t+1}) = (1 + \kappa_{d,1} A_{d,d})^2 \varphi_d^2 \sigma_t^2 + \sum_{j=2,3,4} \Lambda_d' h_j h_j' \Lambda_d q_t, \tag{2.7}$$

$$VRP_t = \left(\sum_{j=1,5} \Lambda_d' h_j h_j' \Lambda_d s_{q,1} + \sum_{j=2,3,4} \Lambda_d' h_j h_j' \Lambda_d s_{q,2} \right) q_t, \tag{2.8}$$

$$dp_t = -A_{0,d} - A_{d,x} x_t - A_{d,\sigma} \sigma_t^2 - A_{d,q} q_t - A_{d,d} \Delta d_t, \tag{2.9}$$

where $s_{q,1} = -(\varphi_x s_{\sigma,x} h_2' + h_3')\Lambda$ and $s_{q,2} = -(\varphi_x s_{q,x} h_2' + s_{q,\sigma} h_3' + \varphi_q h_4')\Lambda$. We will impose these "structural" restrictions on the empirical model estimated below.

Even though our empirical strategy of substituting out consumption means that some of the parameters in the autoregressive loading matrix F and the volatility loading matrix H are not identified, the specific structures for the two loading matrices still provide useful guidance on how to restrict the dynamics. In particular, denote

[9] As further detailed in the supplementary Appendix, the market prices of risks also depend implicitly on the coefficients in the wealth-consumption ratio $\nu_t = A_0 + [0, A_x, A_\sigma, A_q, 0]' Y_t$ and the price-dividend ratio $w_t \equiv -dp_t = A_{d,0} + [0, A_{d,x}, A_{d,\sigma}, A_{d,q}, A_{d,d}]' Y_t$.

the sub-vector of Y_t that excludes consumption growth by $f_t \equiv [\sigma_t^2, q_t, \Delta d_t, x_t]'$, it follows that

$$f_{t+1} = +f_t + S\epsilon_{t+1}, \tag{2.10}$$

where

$$= \begin{pmatrix} \rho_\sigma & 0 & 0 & 0 \\ 0 & \rho_q & 0 & 0 \\ 0 & 0 & \rho_d & \phi_{dx} \\ 0 & 0 & 0 & \rho_x \end{pmatrix} \qquad S = \begin{pmatrix} 1 & 0 & 0 & s_{\sigma,x} \\ s_{q,\sigma} & 1 & 0 & s_{q,x} \\ s_{d,\sigma} & s_{d,q} & 1 & s_{d,x} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.11}$$

and the vector of innovations $\epsilon_{t+1} \equiv [\sqrt{q_t} z_{\sigma,t+1}, \varphi_q \sqrt{q_t} z_{q,t+1}, \varphi_d \sigma_t z_{d,t+1}, \varphi_x \sqrt{q_t} z_{x,t+1}]'$ is conditionally heteroskedastic.[10]

2.3 "Structural" Estimation Results

The consumption-based asset pricing model with volatility uncertainty, outlined in the previous section, imposes a number of restrictions pertaining to the dynamic dependencies and possible feedback effects between the expected variance, the variance risk premium, the dividend growth rate, and the dividend-price ratio. Our new "structural" factor GARCH model is designed to honor these restrictions within a tractable econometric framework.

2.3.1 Data Description

Our empirical investigations are based on end-of-month S&P 500 index returns, as a proxy for the aggregate market portfolio, and the S&P 500 dividend payments, as a proxy for the corresponding aggregate cash flows. All of our S&P 500 data are obtained from DataStream, and cover the period from January 1990 to November

[10] The value of is immaterial to all of our predictability results. Also, the reordering of the elements in f_t relative to Y_t merely serves to facilitate comparisons with other benchmark models below, and does not affect any of the results.

2011, for a total of 262 monthly observations.[11]

Following standard practice in the literature, we use the trailing 12-month dividend-price ratio to account for the strong seasonality inherent in the dividend payouts; see, e.g., the discussion in Bollerslev and Hodrick (1995). Accordingly, the month t log dividend-price ratio dp_t, is defined by,

$$dp_t = \log\left(\frac{Div_{t-11} + ... + Div_t}{12P_t}\right), \tag{2.12}$$

where Div_t denotes the dividend payments from the end-of-month $t-1$ to the end-of-month t, and P_t denotes the end-of-month t price. Our measures for the month $t+1$ log dividend growth rate Δd_{t+1} and the log returns including dividends $r_{t,t+1}$, are similarly defined from this ratio as,

$$\Delta d_{t+1} = \log\left(\frac{Div_{t-10} + ... + Div_{t+1}}{Div_{t-11} + ... + Div_t}\right), \tag{2.13}$$

$$r_{t,t+1} = \log\left(\frac{P_{t+1} + \frac{Div_{t-10}+...+Div_{t+1}}{12}}{P_t}\right), \tag{2.14}$$

with longer-run dividend growth rates and multi-period returns obtained by summation.

We consider three distinct empirical variation measures: the options implied variation IV_t, the expected return variation ERV_t, and the variance risk premium VRP_t. Our measure for the options implied variation is based on the square of the Chicago Board of Options Exchange (CBOE) VIX volatility index. This model-free measure is (approximately) equal to the market risk-neutral, or $\mathbb{Q}$, expectation of the one-month-ahead return variation under very general assumptions. Our construction of the corresponding actual, or $\mathbb{P}$, expectation, is based on the linear projection of the monthly realized variance $RV_{t,t+1}$ on its lagged daily $RV_{t-\frac{1}{22},t}$, weekly $RV_{t-\frac{5}{22},t}$,

[11] While the S&P 500 data are obviously available over a much longer sample period, some of the key variation measures employed in our analysis are only available starting in 1990.

and monthly $RV_{t-1,t}$ values, along with the implied variation IV_t; i.e.,[12]

$$ERV_{t,t+1} \;=\; \alpha_0 + \alpha_1 RV_{t-\frac{1}{22},t} + \alpha_2 RV_{t-\frac{5}{22},t} + \alpha_3 RV_{t-1,t} + \alpha_4 IV_t. \tag{2.15}$$

This mimics the popular HAR-RV model proposed by Corsi (2009b). Importantly, the addition of IV_t as an additional right-hand-side variable imbues the formulation in (2.15) with an additional persistent long-run predictor variable, which in the traditional HAR-RV model would be captured by longer-run realized variation measures.[13] Finally, our measure for the variance risk premium is simply given by the difference between our risk-neutral and statistical expectations of the one-month-ahead return variation; i.e., $VRP_t = IV_t \;-\; ERV_t$.

To illustrate the basic features of the different variables, Figure 2.1 plots the monthly time series of stock returns, dividend growth rates, dividend-price ratios, and variance risk premia. The large losses in market values and the increased volatility during the recent economic downturn are immediately evident in the plots of the returns and cash flows. The plot for the dividend-yields shows a sharp drop throughout the 1990s, but an increase after the burst of the tech bubble in 2001, reaching a new peak in the fourth quarter of 2008 around the advent of the global financial crisis and the stock market crash.[14] The variance risk premium shown in the last panel is on average positive with occasional negative spikes, the largest of which oc-

[12] Our regression-based estimates of the α's rely on overlapping daily observations for all of the variation measures, thus implicitly assuming that the same relationship holds every day of the month. This greatly enhances the accuracy of the estimates compared to the estimates obtained by the use of non-overlapping monthly observations only.

[13] We also experimented with decomposing the realized variation measures into their continuous and discontinuous parts. Although this often helps for shorter-run forecasting, consistent with the results in Andersen, Bollerslev, and Diebold (2007), we found that the monthly forecasts and R^2s from these more elaborate models were virtually the same as the ones from the simple-to-implement HAR-RV type formulation in (2.15).

[14] The sharp decline observed in the 1990s has been attributed to firms' substitution of dividend payments by share repurchases; see, e.g., Koijen and van Nieuwerburgh (2011), along with the earlier related discussion in Bagwell and Shoven (1989).

cur in the fall of 2008 at the onset of the financial crises. Summary statistics for the same four variables, along with the options implied and expected variation measures underlying the variance risk premium, are reported in Table 2.1.

We turn next to our new present value framework and "structural" model designed to describe these general features and inherent dynamic dependencies.

2.3.2 "Structural" Factor GARCH

The dynamics of the asset pricing model in Section 2.2 is succinctly summarized by the state vector f_t and equations (2.10) and (2.11). The state vector f_t is, of course, not directly observable. To circumvent this, we define the "observable" state vector $X_t \equiv [ERV_t, VRP_t, \Delta d_t, dp_t]'$. From the solution of the model, the X_t vector is directly related to the latent f_t vector by the linear equations,[15]

$$X_t = \mu_X + Q f_t \qquad\qquad Q = \begin{pmatrix} Q_{1,1} & Q_{1,2} & 0 & 0 \\ 0 & Q_{2,2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -A_{d,\sigma} & -A_{d,q} & -A_{d,d} & -A_{d,x} \end{pmatrix}, \tag{2.16}$$

where $Q_{1,1} = (1 + \kappa_{d,1}A_{d,d})^2\varphi_d^2\rho_\sigma$, $Q_{1,2} = \sum_{j=2,3,4}\Lambda_d h_j h_j'\Lambda_d\rho_q$, and $Q_{2,2} = (1 + \kappa_{d,1}A_{d,d})^2 s_{q,1} + \sum_{j=2,3,4}\Lambda_d' h_j h_j'\Lambda_d s_{q,2}$. Given the standard set of assumptions about the structural parameter values typically employed in the long-run risk literature, all of the Q parameters would be positive. Conversely, $A_{d,\sigma}$, $A_{d,q}$, and $A_{d,d}$ would all be negative, while $A_{d,x}$ is naturally expected to be positive.

Now combining the model for f_t in equations (2.10) and (2.11) with the expression for X_t in equation (2.16), it follows that

$$BX_{t+1} = \tilde{\mu} + \tilde{B}X_t + \tilde{S}\tilde{\epsilon}_{t+1}, \qquad \tilde{\epsilon}_{t+1} = \tilde{G}_t z_{t+1}, \tag{2.17}$$

[15] Additional details concerning the solution of the model is available in the supplementary Appendix.

where $\tilde{G}_t = \text{diag}[Q_{1,1}\sqrt{q_t}, Q_{2,2}\varphi_q\sqrt{q_t}, \varphi_d\sigma_t, -A_{d,x}\varphi_x\sqrt{q_t}]$, and[16]

$$B = \begin{pmatrix} 1 & -\frac{Q_{1,2}}{Q_{2,2}} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{A_{d,\sigma}}{Q_{1,1}} & \frac{Q_{1,1}A_{d,q}-A_{d,\sigma}Q_{1,2}}{Q_{1,1}Q_{2,2}} & \frac{\rho_d}{1-\kappa_{d,1}\rho_d} & 1 \end{pmatrix} \qquad \tilde{\rho} = \begin{pmatrix} \rho_\sigma & 0 & 0 & 0 \\ 0 & \rho_q & 0 & 0 \\ 0 & 0 & \rho_d & \frac{\phi_{dx}}{-A_{d,x}} \\ 0 & 0 & 0 & \rho_x \end{pmatrix} \tag{2.18}$$

$$\tilde{S} = \begin{pmatrix} 1 & 0 & 0 & \frac{Q_{1,1}}{-A_{d,x}}s_{\sigma,x} \\ \frac{Q_{2,2}}{Q_{1,1}}s_{q,\sigma} & 1 & 0 & \frac{Q_{2,2}}{-A_{d,x}}s_{q,x} \\ \frac{1}{Q_{1,1}}s_{d,\sigma} & \frac{1}{Q_{2,2}}s_{d,q} & 1 & \frac{1}{-A_{d,x}}s_{d,x} \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.19}$$

Multiplying the "structural" VAR in equation (2.17) by B^{-1}, the corresponding reduced form VAR(1) representation for X_{t+1} becomes,

$$X_{t+1} = B^{-1}\tilde{\mu} + \Phi X_t + u_{t+1}, \tag{2.20}$$

where Φ=$B^{-1}$$B$, $u_{t+1} = \Phi_0^{-1}\tilde{\epsilon}_{t+1}$, and $\Phi_0^{-1} = B^{-1}\tilde{S}$. As this representation makes clear, ignoring the heteroskedasticity in the reduced form shocks u_{t+1}, and interpreting the model for X_{t+1} in (2.17) as a standard homoskedastic VAR(1), the B and $\tilde{S}$ matrices would not be jointly identified. In empirical macroeconomics, this lack of identification is usually "solved" by imposing that Φ_0 is lower triangular. However, as argued by Sentana and Fiorentini (2001), Rigobon (2003) and Rigobon and Sack (2003), among others, under the maintained assumption that the underlying "structural" shocks are independent, it is possible to identify the Φ_0 matrix, and in turn both B and $\tilde{S}$, through the heteroskedasticity in $\tilde{\epsilon}_{t+1}$.

Meanwhile, rather than specifying the time-varying covariance matrix for the "structural" shocks to be an explicit function of the latent q_t and σ_t^2 risk factors, in the implementation reported on below we adopt a more flexible and empirically realistic

[16] As explained in more detail in the supplementary Appendix, the matrix B matrix is obtained from the matrix Q by normalizing its diagonal elements to unity.

GARCH approach for characterizing the dynamic dependencies in $\tilde{\epsilon}_{t+1}$. Specifically, let Σ_{t+1} denote the conditional covariance matrix of $\tilde{\epsilon}_{t+1}$. We will then assume that Σ_{t+1} may be described by the following relatively simple yet flexible diagonal GARCH(1,1) model,

$$\text{diag}(\Sigma_{t+1}) = (I - \Gamma - \Upsilon)\Theta_0^{-1}\varpi_u + \Gamma\text{diag}(\Sigma_t) + \Upsilon\tilde{\epsilon}_t^2, \tag{2.21}$$

where $\Theta_0 = \Phi_0^{-1} \odot \Phi_0^{-1}$, and ϖ_u denotes the unconditional covariance matrix of the reduced form shocks $u_{t+1} = \Phi_0^{-1}\tilde{\epsilon}_{t+1}$. Consequently, the second order dynamics of u_{t+1} will follow the more complicated non-diagonal GARCH(1,1) structure,[17]

$$\text{vec}(\Omega_{t+1}) = \Theta_1(I - \Gamma - \Upsilon)\Theta_0^{-1}\varpi_u + \Theta_1\Gamma\Theta_0^{-1}\text{diag}(\Omega_t) + \Theta_1\Upsilon\Theta_2\text{vec}(u_t u_t'). \tag{2.22}$$

By explicitly parameterizing this implied conditional heteroskedasticity in u_{t+1}, it is possible to identify and separately estimate *all* of the "structural" parameters in (2.17)-(2.19).

The diagonal GARCH(1,1) model in (2.21) freely parametrizes the persistence in the "structural" shocks. Consistent with our initial estimates of the model, and the implication from the underlying consumption-based asset pricing model, we impose the restriction that the autoregressive dependencies in the GARCH expected variance and the dividend-price ratio are the same, i.e., $\Gamma_{1,1} + \Upsilon_{1,1} = \Gamma_{4,4} + \Upsilon_{4,4} = \rho_q$. Guided by our initial diagnostic tests, we also restrict the dividend growth shock to have only ARCH and no GARCH effect, i.e., $\Gamma_{3,3} = 0$. All-in-all, this leaves us with a total of nine conditional variance parameters to be estimated.

Let ξ denote the vector of stacked parameters comprised of the conditional mean parameters in B, $\tilde{S}$, $\tilde{\mu}$, and $\tilde{}$, along with the conditional variance parameters in Γ, Υ, and ϖ_h. Assuming that the reduced form shocks u_{t+1} are jointly normally

[17] More formally, $\Theta_1 = (\Phi_0^{-1}\Phi_0^{-1})I_l$, $\Theta_2 = [\text{vec}(\Phi_{0,(1)}^{-1'}\Phi_{0,(1)}^{-1}), \text{vec}(\Phi_{0,(2)}^{-1'}\Phi_{0,(2)}^{-1}), \text{vec}(\Phi_{0,(3)}^{-1'}\Phi_{0,(3)}^{-1}), \text{vec}(\Phi_{0,(4)}^{-1'}\Phi_{0,(4)}^{-1})]'$, where $\Phi_{0,(i)}^{-1}$ denotes the i^{th} row of the square matrix Φ_0^{-1}, and the 16×4 matrix I_l helps to transform the vector $\text{vec}(\Omega_t)$ into diagonal matrix form.

distributed, the logarithm of the density for X_{t+1} conditional on X_t and Ω_{t+1}, or equivalently the contribution to the log-likelihood function coming from X_{t+1}, may be expressed as,

$$\begin{aligned} L_t(X_{t+1},\xi) = &-2\log 2\pi - \frac{1}{2}\log|\Omega_t| \\ &- \frac{1}{2}(X_{t+1} - B^{-1}\tilde{\mu} - \Phi X_t)'\Omega_t^{-1}(X_{t+1} - B^{-1}\tilde{\mu} - \Phi X_t) \qquad (2.23) \\ = &-2\log 2\pi - \frac{1}{2}\log|\Sigma_t| + \log|\tilde{S}^{-1}B| \\ &- \frac{1}{2}(X_{t+1} - B^{-1}\tilde{\mu} - B^{-1}BX_t)'\Xi(X_{t+1} - B^{-1}\tilde{\mu} - B^{-1}BX_t). \qquad (2.24) \end{aligned}$$

where $\Xi = \tilde{S}^{-1}B\Sigma_t^{-1}B'\tilde{S}^{-1'}$. Even if the assumption of conditional normality is violated empirically, the estimate for ξ obtained by maximizing the resulting log-likelihood function, defined by summing (2.23) over the full sample, remains consist and asymptotically normally distributed under quite general conditions; see, e.g., Bollerslev and Wooldridge (1992).

The long-run implications from multivariate GARCH models can be very sensitive to estimation errors and small perturbations in a few parameters. To help guard against this, we augment the Gaussian-based score for the "structural" VAR-GARCH model with an additional set of moment conditions designed to ensure that the unconditional variances of the reduced form errors implied by the model match their standard VAR-based analogues.[18] Expressing this additional set of moments in parallel to equation (2.23) and the contribution to the likelihood function coming

[18] This mirrors the variance targeting approach originally advocated by Engle and Mezrich (1996). However, in contrast to that two-step approach, the GMM-based procedure applied here jointly estimates all of the parameters in ξ in a single step.

from X_{t+1}, we have

$$W_t(X_{t+1},\xi) \;=\; \varpi_u - \text{diag}\left((X_{t+1}-\mu^{\text{OLS}}-\Phi^{\text{OLS}}X_t)(X_{t+1}-\mu^{\text{OLS}}-\Phi^{\text{OLS}}X_t)'\right), \tag{2.25}$$

where the "OLS" superscript indicates the parameters obtained from equation-by-equation least squares estimation of the reduced form VAR. The estimates for ξ reported below are obtained by applying standard iterated GMM to the conditional set of moments defined by the score for the conditional density in (2.23), say $\partial_\xi L_t(X_{t+1},\xi)$, augmented with the moment conditions in (2.25),[19]

$$g(X_{t+1},\xi) = \begin{pmatrix} \partial_\xi L_t(X_{t+1},\xi) \\ W_t(X_{t+1},\xi) \end{pmatrix}. \tag{2.26}$$

We turn next to a discussion of the resulting $\hat{\xi}$, and the implications of the estimates in regards to the dynamics of the systematic risk factors and the dependencies among the "structural" shocks.

2.3.3 Estimation Results

The dynamic dependencies in the observable state vector $X_t = [ERV_t, VRP_t, \Delta d_t, dp_t]'$ underlying our GMM estimation is directly related to the latent state vector $f_t = [\sigma_t^2, q_t, \Delta d_t, x_t]'$ of interest by the affine equation $X_t = \mu_X + Qf_t$. This allows us to infer both the contemporaneous interaction matrix Q and the autoregressive matrix describing the mean dynamics in $f_{t+1} = +f_t + S\epsilon_{t+1}$ from the estimates for B and ˜ based on $BX_{t+1} = \tilde{\mu} + \tilde{}BX_t + \tilde{S}\tilde{\epsilon}_{t+1}$, and the relations in equation (2.18) above. Similarly, the estimated volatility loading matrix $\tilde{S}$ for the observable state vector X_t allow us to infer the volatility loading matrix S for the latent state vector f_t from equation (2.19), while the estimated volatility dynamics of the $\tilde{\epsilon}_{t+1}$ shocks effectively

[19] This idea of augmenting the likelihood function with additional information mirrors the use of quasi-Bayesian priors, applied in a different context by, e.g., Hamilton (1991), and may also be seen as a form of shrinkage type estimation.

determines the implied volatility dynamics of the "structural" ϵ_{t+1} shocks.

We begin with a discussion of the estimates for B and $\tilde{}$,

$$\hat{B} = \begin{pmatrix} 1 & \underset{(0.11)}{-0.02} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \underset{(0.03)}{-0.60} & \underset{(0.10)}{-1.44} & -0.19 & 1 \end{pmatrix} \qquad \hat{\tilde{}} = \begin{pmatrix} \underset{(0.05)}{0.64} & 0 & 0 & 0 \\ 0 & \underset{(0.07)}{0.46} & 0 & 0 \\ 0 & 0 & \underset{(0.03)}{-0.23} & \underset{(0.004)}{-0.002} \\ 0 & 0 & 0 & \underset{(0.009)}{0.988} \end{pmatrix} \quad (2.27)$$

where the numbers in parentheses represent asymptotic standard errors. With the exception of $B_{1,2}$ and $\tilde{}_{3,4}$, all of the individual parameter estimates are highly statistically significant. All of the estimates also have the "correct" signs vis-a-vis the implications from the equilibrium-based model and the "structural" VAR.

In particular, the negative estimates for the loadings for the dividend price ratio reported in the last row of the B matrix are consistent with the idea that the two volatility components σ_t^2 and q_t, and cash flow growth Δd_t, are all genuine risk factors with negative market prices of risks.[20] Within the context of the standard Bansal and Yaron (2004a) long-run risk model, these negative contemporaneous relationships between the dividend-price ratio and the other state variables, or risk factors, are critically dependent on the risk aversion parameter $\gamma > 1$ and the intertemporal elasticity of substitution $\psi > 1$. As such, our "structural" estimation results indirectly support this commonly invoked set of assumptions.

Our estimate for $\tilde{}_{4,4} =_x= 0.988$ also points to a highly persistent and very accurately estimated long-run risk factor. This contrasts with the typical practice of simply fixing the long-run persistence coefficient at some "large" value, as in, e.g., Bansal, Gallant, and Tauchen (2007b), and clearly highlights the advantages of the more structured GMM estimation approach and richer data sources applied here.

[20] Note that the market price of dividend risk $B_{4,3} = -0.19$ is imputed to by the constraint $A_{d,d} = \frac{\rho_d}{1-\kappa_{d,1}\rho_d}$ imposed in equation (2.18).

Meanwhile, even though our estimate for $\phi_{dx} = \tilde{}_{3,4} = \frac{\phi_{dx}}{-A_{d,x}} = -0.002$ is "correctly" signed, the parameter is not significantly different from zero, and as such offers only limited support to the idea that the long-run risk factor x_t contemporaneously impacts cash flows Δd_t.

Interestingly, our use of more accurate volatility measures results in a much more persistent consumption variance estimate $\tilde{}_{1,1} =_{\sigma}= 0.64$ compared to the estimates recently reported in Campbell, Giglio, Polk, and Turley (2013). Moreover, our estimates for $\tilde{}_{1,1} =_{\sigma}= 0.64 > \tilde{}_{2,2} =_{q}= 0.46$ imply that the consumption variance σ_t^2 is more persistent than the variance-of-variance q_t, or economic uncertainty, which is directly in line with the implicit assumptions invoked in the calibrations reported in Bollerslev, Tauchen, and Zhou (2009).

Turning to our estimates for the volatility dependence matrix $\tilde{S}$,

$$\hat{\tilde{S}} = \begin{pmatrix} 1 & 0 & 0 & \underset{(0.04)}{0.08} \\ \underset{(0.06)}{-0.29} & 1 & 0 & \underset{(0.02)}{-0.09} \\ \underset{(0.05)}{-0.36} & \underset{(0.08)}{-0.09} & 1 & \underset{(0.03)}{0.15} \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.28}$$

all of the individual parameters, except $\tilde{S}_{3,2}$, are again highly statistically significant. This clearly underscores the idea that multiple volatility factors are indeed needed to accurately describe the dynamic dependencies observed in the data, and that the standard long-run risk model with a single stochastic volatility factor is misspecified. To more fully appreciate this and the other implications of the estimates recall again the relationship between $\tilde{S}$ and the "structural" S matrix for the latent state vector in equation (2.19).

It follows from this relation that shocks to cash flow growth Δd_t are adversely affected by shocks to the long-run risk component x_t, as $s_{d,x} \propto -\tilde{S}_{3,4} = -0.15$.[21]

[21] We use the symbol $\propto$ to denote proportional to.

This is consistent with the idea that companies tend to distribute more in dividends when long-run growth opportunities are poor. The "structural" long-run risk shock affects the two variance processes σ_t^2 and q_t in opposite directions. Good news about long-run consumption growth reduces the consumption variance, as $s_{\sigma,x} \propto -\tilde{S}_{1,4} = -0.08 < 0$, but increases economic uncertainty, as $s_{q,x} \propto -\tilde{S}_{2,4} = 0.09 > 0$. The first effect represents the well known "leverage effect", whereby a negative growth shock is associated with higher volatility, and vice versa. The second effect, however, is more subtle. Since q_t directly affects the time-varying volatility of the long-run risk component, a positive $s_{q,x}$ implies that when a positive $z_{x,t}$ shock occurs, the volatility of next period's $\epsilon_{x,t+1}$ will also be higher, and vice versa. Intuitively, this could happen when good news in consumption growth is accompanied by better investment opportunities, in turn resulting in higher economic uncertainty, possibly due to over-investment. Interestingly, our estimates for $\tilde{S}$ also suggest that $s_{q,\sigma} \propto \tilde{S}_{2,1} = -0.29 < 0$, implying that a positive "structural" shock to consumption volatility σ_t^2 reduces the uncertainty of volatility q_t. This effect is naturally interpreted as a new "leverage effect" between volatility and volatility-of-volatility.[22]

Our identification and estimation of the "structural" model parameters rely crucially on the presence of time-varying conditional heteroscedasticity in the ϵ_{t+1} shocks. The GMM parameter estimates for the "structural" factor GARCH model describing this heteroscedasticity are reported in Table 2.2. As the table shows, all of the shocks do indeed exhibit highly significant (G)ARCH effects.[23] The overall good fit of the model is also supported by the conventional J-test statistic for general model

[22] This new equilibrium-based "leverage effect" is also consistent with the asymmetries in daily and high-frequency intraday VIX and S&P 500 returns documented in Aboura and Wagner (2012) and Bollerslev, Osterrieder, Sizova, and Tauchen (2013), respectively.

[23] The significance of the (G)ARCH effects is also indirectly supported by Ljung-Box tests for residual serial correlation in the raw and standardized absolute residuals from the model; further details concerning these results are available upon request. This, of course, is directly in line with the burgeoning literature on the estimation of reduced form GARCH and stochastic volatility models for a wide array of other financial and macroeconomic time series.

misspecification and the minimized value of the GMM objective function equal to 12.76, which corresponds to a p-value of 0.12 in the relevant asymptotic chi-square distribution.[24]

In order to further gauge the quality of the fit afforded by the model, Figure 2.2 plots the time-series of "structural" shocks associated with each of the four equations. The top two panels show the volatility shocks $z_{\sigma,t}$ and $z_{q,t}$. Both of these shocks experienced unprecedented large, albeit opposite signed, realizations during the 2007-2009 "Great Recession." Interestingly, neither one of the earlier 1990-1991 and 2001-2002 NBER-dated recessions were accompanied by especially large "structural" volatility shocks. The general time-series pattern of the equilibrium-based cash flow shocks $z_{\Delta d,t}$ appear quite similar to that of the normalized cash flow news in Campbell, Giglio, Polk, and Turley (2013). Although not quite as dramatic as for the two volatility shocks, the permanent growth shocks $z_{x,t}$ also experienced their most extreme realizations during the "Great Recession." This basic dynamic pattern in the equilibrium-based growth shocks is again quite similar to that of the normalized discount rate news shocks reported in Campbell, Giglio, Polk, and Turley (2013).[25]

In lieu of these findings and generally supportive diagnostic tests for the "structural" factor GARCH model, we turn next to our main empirical investigations, showing how incorporating the additional variance-related state variables in the equilibrium-based model help shed new light on the return and dividend growth predictability patterns inherent in the data.

[24] By contrast, the two alternative specifications discussed in the supplementary Appendix, one closer to Drechsler and Yaron (2011b) with $G_t = \text{diag}[\sigma_t, \sqrt{q_t}, \sigma_t, \sqrt{q_t}, \sigma_t]$, and one closer to Branger and Vòlkert (2012) with $G_t = \text{diag}[\sigma_t, \sigma_t, \sigma_t, \sqrt{q_t}, \sigma_t]$, result in GMM-based J-statistics equal to 26.31 and 37.02, respectively, with corresponding p-values essentially zero.

[25] This is also consistent with the findings in Lettau and Ludvigson (2013), who suggest that large negative permanent growth shocks might have adversely affected housing wealth.

2.4 Model Implied Return and Cash Flow Predictability

Our predictability analysis is based on recasting the "structural" factor GARCH model in the form of an expanded VAR system, along with the use of the standard Campbell-Shiller approximation for expressing the return as a function of the observable state variables.

2.4.1 VAR and Predictability

The first order VAR for the state vector X_t=$[ERV_t, VRP_t, \Delta d_t, dp_t]$ implied by the "structural" factor GARCH model in equation (2.20) doesn't directly involve the return. However, by the standard Campbell-Shiller approximation, the return may be conveniently expressed as $r_{t,t+1} = \kappa_{d,0} - \kappa_{d,1} dp_{t+1} + dp_t + \Delta d_{t+1}$.[26] Combining this equation for $r_{t,t+1}$ with the VAR for X_{t+1}, it follows that

$$r_{t,t+1} \;=\; \mu_r + (l_1\Phi \;+\; e_4)X_t + l_1\Phi_0^{-1}\tilde{\epsilon}_{t+1}, \tag{2.29}$$

where μ_r collects all of the relevant constant terms, $l_1 \equiv (0, 0, 1, -k_{1,d})$, and the selection vector $e_4 \equiv (0, 0, 0, 1)$. Iterating the VAR for X_t forward, it is therefore possible to derive closed-form expressions for the model-implied multi-period return $r_{t,t+h} = r_{t,t+1} + ... + r_{t+h-1,t+h}$ regressions based any explanatory variable spanned by the X_t state vector.

In the analysis reported on below we will focus on the three key predictor variables: the log dividend-price ratio dp_t, the variance risk premium VRP_t, and the expected variation ERV_t. In particular, consider the regression of the h-period re-

[26] The accuracy of the Campbell-Shiller approximation has recently been corroborated by Engsted, Pedersen, and Tanggaard (2012). By definition $\kappa_{d,1} = \exp(-E(dp_t))[1 + \exp(-E(dp_t))]^{-1}$. In the estimation results reported on below we rely on the sample average of the monthly dividend-price ratio from January 1965 to November 2011 when calculating $E(dp_t)$, implying a value of $\kappa_{d,1} = 0.9976$.

turns on the dividend-price ratio,

$$\frac{1}{h}\sum_{i=1}^{h} r_{t,t+i} \;=\; \alpha_{r,dp} \;+\; \beta_{r,dp}(h)\cdot dp_t \;+\; \varsigma_{t,t+h}. \tag{2.30}$$

By similar arguments to the ones in Hodrick (1992) and Campbell (2001), it is possible to show that

$$\beta_{r,dp}(h) = \frac{(l_1\Phi + e_4)(I-\Phi)^{-1}(I-\Phi^h)C(0)e_4'}{e_4 C(0) e_4'} \tag{2.31}$$

where $C(0) = \sum_{j=0}^{\infty}\Phi^j\Phi_0^{-1}\text{diag}(\Theta_0^{-1}\varpi_u)\Phi_0^{-1'}\Phi^{j'}$ denotes the model-implied unconditional covariance matrix for X_t, and $e_4 \equiv (0,0,0,1)$.[27] Similarly, the implied coefficients for the return predictability regressions based on VRP_t and ERV_t may be expressed in close form as,

$$\beta_{r,VRP}(h) = \frac{(l_1\Phi + e_4)(I-\Phi)^{-1}(I-\Phi^h)C(0)e_2'}{e_2 C(0) e_2'} \tag{2.32}$$

$$\beta_{r,ERV}(h) = \frac{(l_1\Phi + e_4)(I-\Phi)^{-1}(I-\Phi^h)C(0)e_1'}{e_1 C(0) e_1'} \tag{2.33}$$

where the e_1 and e_2 selection vectors are defined in an obvious manner.[28]

In parallel to equation (2.29) for the returns, the growth rate dynamics implied by the "structural" factor GARCH may be expressed in linear form as,

$$\Delta d_{t+1} = \mu_d + e_3\Phi X_t + e_3\Phi_0^{-1}\tilde{\epsilon}_{t+1}, \tag{2.34}$$

where μ_d collects all the relevant constant terms. Thus, replacing $l_1\Phi + e_4$ with $e_3\Phi$

[27] In the empirical results reported on below, we truncate the infinite sum in the expression for $C(0)$ at 120, or ten years; see Bollerslev and Hodrick (1995) for further discussion along these lines.

[28] Analytical expressions for the R^2s from the regressions may be derived in a similar manner. Specifically, for the dividend-price ratio regression $R^2_{r,dp}(h) = h^2\beta^2_{r,dp}(h)e_4C(0)e_4'/Var(\sum_{j=1}^{h} r_{t,t+j})$, where $Var(\sum_{j=1}^{h} r_{t,t+j}) = h(l_1\Phi + e_3)C(0)(l_1\Phi + e_3)' + hl_1(I-\Phi)C(0)l_1' + \sum_{i=1}^{h-1} 2(h-i)((l_1\Phi + e_3)\Phi^i C(0)(l_1\Phi + e_3)' + (l_1\Phi + e_3)\Phi^{i-1}(I-\Phi)C(0)e_3')$.

in the formulas for the regression coefficients above, comparable expressions for the cash flow predictability regression coefficients $\beta_{\Delta d,dp}(h)$, $\beta_{\Delta d,VRP}(h)$, and $\beta_{\Delta d,ERV}(h)$ are readily available. When interpreting these coefficients, it is important to keep in mind the relationship $E_t(\Delta d_{t+1}) = \phi_{dx} x_t + \rho_d \Delta d_t$ implied by equations (2.10) and (2.11), and the fact that within the "structural" model the expected value of next periods dividend growth rate is linearly related to the lagged dividend growth rate and the long-run risk component.

2.4.2 Model-Implied Reduced Form VAR Estimates

The reduced form VAR parameter matrix Φ and the unconditional covariance matrix $C(0)$ for X_t entering the expressions for the predictive regression coefficients in equations (2.31)-(2.33) could, of course, be estimated directly by OLS equation-by-equation. However, that obviously would ignore any of the equilibrium-based "structural" restrictions. It also would not permit the separate identification of the contemporaneous Φ_0 matrix entering the expressions for the return and dividend growth rate in equations (2.29) and (2.34), respectively.

Instead, the Φ_0 and Φ parameter matrices may both be deduced from the "structural" factor GARCH model parameters and the relations $\Phi = B^{-1}\tilde{}B$ and $\Phi_0^{-1} = B^{-1}\tilde{S}$ derived above. Substituting the previously discussed estimates for $B, \tilde{}$ and $\tilde{S}$ into these expressions, yields,

$$\hat{\Phi} = \begin{pmatrix} \underset{(0.05)}{0.64} & \underset{(0.020)}{-0.003} & 0 & 0 \\ 0 & \underset{(0.07)}{0.46} & 0 & 0 \\ \underset{(0.002)}{0.001} & \underset{(0.005)}{0.002} & \underset{(0.03)}{-0.23} & \underset{(0.004)}{-0.002} \\ \underset{(0.03)}{-0.21} & \underset{(0.13)}{-0.76} & \underset{(0.03)}{-0.23} & \underset{(0.087)}{0.988} \end{pmatrix} \qquad \hat{\Phi}_0 = \begin{pmatrix} \underset{(0.033)}{0.995} & \underset{(0.11)}{0.02} & 0 & \underset{(0.04)}{0.08} \\ \underset{(0.06)}{-0.29} & 1 & 0 & \underset{(0.02)}{-0.09} \\ \underset{(0.06)}{-0.34} & \underset{(0.08)}{-0.09} & 1 & \underset{(0.03)}{0.15} \\ \underset{(0.09)}{0.11} & \underset{(0.13)}{1.44} & \underset{(0.02)}{0.19} & \underset{(0.03)}{0.94} \end{pmatrix} \tag{2.35}$$

where the numbers in parentheses represent standard errors derived by the delta-

method.

Based on these estimates for Φ and Φ_0, the return equation in (2.29) may be expressed numerically as,

$$\begin{aligned} r_{t,t+1} = \underset{(0.03)}{0.05} + \underset{(0.03)}{0.20} ERV_t + \underset{(0.12)}{0.76} VRP_t - \underset{(0.002)}{0.0013} \Delta d_t + \underset{(0.011)}{0.013} dp_t \\ \underset{(0.08)}{-0.47} \tilde{\epsilon}_{\sigma,t+1} \underset{(0.11)}{-1.52} \tilde{\epsilon}_{q,t+1} + \underset{(0.04)}{0.81} \tilde{\epsilon}_{\Delta d,t+1} \underset{(0.03)}{-0.79} \tilde{\epsilon}_{x,t+1}. \end{aligned} \tag{2.36}$$

Of course, this "estimated" return equation does not actually rely on the return data, but instead is deduced from our estimates for the equilibrium-based model and the observable state vector involving the dividend growth rate and the log dividend-price ratio. Again, this mirrors the approach of Cochrane (2008). However, in contrast to the return equation therein, which only involves the dividend-price ratio, we purposely include the two variance variables, both of which enters with highly significant coefficients.

Further underscoring the importance of incorporating the variation measures into the analysis, the model-implied loadings for all of the "structural" shocks are also highly significant. Among the four shocks, the ones for the long-run risk component and the consumption variance uncertainty have the largest impacts, accounting for 43 percent ($z_{x,t}$) and 26 percent ($z_{q,t}$) of the unexpected unconditional return variation, respectively. The "estimated" return equation in (2.36) also implies that the total one-month explainable return variation equals 9 percent, far exceeding that afforded by traditional univariate return predictability regressions that does not include ERV_t and VRP_t.

Explicitly writing out the second equation for the variance risk premium in the model-implied VAR,

$$VRP_{t+1} = \underset{(0.001)}{0.001} + \underset{(0.07)}{0.46} VRP_t \underset{(0.06)}{-0.29} \tilde{\epsilon}_{\sigma,t+1} + \tilde{\epsilon}_{q,t+1} \underset{(0.02)}{-0.09} \tilde{\epsilon}_{x,t+1}. \tag{2.37}$$

shows that the only "structural" shock that enters the return and VRP equations with the opposite sign is $\tilde{\epsilon}_{q,t}$. Indeed, excluding the impact of the economic uncertainty shock from both equations changes the monthly conditional correlation, or "leverage effect," from a negative -0.09 to a positive 0.66, again reinforcing the importance of jointly modeling *all* of the elements in the X_t state vector.

2.4.3 Model-Implied Predictability Relations

The VAR-based formula for the slope coefficients presented above allow for a direct assessment of the statistical significance of the different predictor variables across different forecast horizons. The formula also allow us to directly assess the enhanced efficiency afforded by the "structural" factor GARCH model compared to the reduced form VAR and simple univariate regression procedures traditionally used in the literature.

To begin, the top panel in Table 2.3 reports the implied slope coefficients for forecasting returns and cash flows by the dividend-price ratio dp_t over long 1- to 10-year horizons, as previously analyzed in the literature. Although the patterns in the estimated coefficients are generally in line with the estimates reported in the existing literature based on longer calendar time spans of data, taken as a whole there is little evidence for any predictability over these long multi-year horizons in the data analyzed here.[29] The results for the shorter within year "structural" and simply unconstrained univariate regressions reported in the lower panel of the table tell a similar story.

The lack of predictability for the long multi-year horizons, is, of course, not too surprising. With only slightly more than twenty years worth of monthly observations

[29] We also experimented with a traditional two-variable homoskedastic VAR for the dividend-price ratio and the dividend growth rate, as in Cochrane (2008), resulting in similar coefficient estimates, but typically larger standard errors, thus highlighting the more accurate inference afforded by explicitly incorporating the equilibrium-based restrictions and the strong heteroskedasticity inherent in the data. Further details concerning these results are available upon request.

any suggestions about statistically significant long-run predictability should be taken with a grain of salt. For the remainder of this section, we will consequently restrict our attention to within-year horizons only.[30]

Turning to our key empirical findings pertaining to the "new" variance related forecasting variables, Figure 2.3 shows the regression slope coefficients for the variance risk premium VRP_t implied by the "structural" factor GARCH model (indicated by dots) along with the corresponding 95 percent confidence intervals (indicated by the shaded area). For comparison purposes, we also include the estimated slope coefficients from simple univariate predictive regressions based on the variance risk premium (indicated by the stars) along with their 95 percent confidence intervals (indicated by the dashed lines). Focusing on the top panel for the returns, both procedures result in significant estimates for up to eight months. It is noteworthy that even though the model-implied point estimates are systematically lower than the unrestricted OLS estimates, they are also less erratic, and the confidence intervals much smaller. Indeed, looking at the numbers in Table 2.4, the t-statistics for testing the null hypothesis of no return predictability are uniformly larger for the "structural" approach.

This discrepancy in the results across the two approaches is even stronger for the cash flow predictability regressions reported in the bottom panel in Figure 2.3. Whereas the estimated slope coefficients from the univariate regressions are all insignificant, the t-statistics associated with the VAR-based model-implied coefficients are all negative and exceed conventional significance levels for up to six months. Hence, not only are higher variance risk premia positively related to future returns, as previously documented in the literature, they also predict lower near-term future

[30] The univariate return regressions reported in Bollerslev, Tauchen, and Zhou (2009) and Drechsler and Yaron (2011b) that in part motivate our analysis also suggest that the return predictability inherent in the variance risk premium is confined to relatively short horizons.

cash flows.[31] This, of course, contrasts with the view commonly expressed in the literature that dividend growth rates are largely unpredictable over short within-year horizons.

Of course, the much-studied classical risk-return trade-off is not based on the variance risk premium, but rather the return variation itself. In spite of the intuitively appealing idea behind such a relationship, empirical attempts at establishing a significant risk-return tradeoff have largely proven futile; see, e.g., the discussion in Bollerslev and Zhou (2006) and Guo and Whitelaw (2006), and the many other references therein. The result for the univariate return regressions based on ERV_t reported in the top panel in Figure 2.4 and Table 2.5 underscore the elusive nature of a simple linear relationship between the expected returns and the expected variation in the data analyzed here. None of the regression coefficients are significant, and most have the "wrong" sign. By contrast, the VAR-based estimates implied by the "structural" model are all positive and marginally significant for return horizons in excess of 4 months.[32]

The difference in the quality of the inference afforded by standard univariate regression-based procedures traditionally employed in the literature and the "structural" approach advocated here is even more dramatic for the cash flow predictions reported in the bottom panel in Figure 2.4. While the simple univariate regressions suggest that the 1-6 months dividend growth rate is unpredictable, the regression coefficients implied by the "structural" model are all highly significant. Interestingly,

[31] This is also related to the observation by Bloom (2009) that an increase in economic uncertainty causes firms to temporarily reduce their investment and hiring, in turn resulting in a short-term productivity drop.

[32] The use of $IV_t = VRP_t + ERV_t$ results in qualitatively similar patterns, but slightly more significant coefficient estimates, compared to the ones reported for ERV_t, thus confirming earlier empirical findings in Bollerslev and Zhou (2006) and Guo and Whitelaw (2006) of a stronger risk-return trade-off when using implied as opposed to realized variation. Still, none of the univariate return regressions based on IV_t result in any significant predictability. Further details of these results are available upon request.

whereas an increase in VRP_t predicts lower future cash flows, and increase in ERV_t is associated with significantly higher future cash flows. Again, this strong empirical evidence for short-run within-year cash flow predictability stands in sharp contrast to the results reported in the existing literature based on other more traditional predictor variables and valuation ratios.

At a more general level, the results for the two different approaches reported in Tables 2.3-2.5 and Figures 2.3-2.4 may also be seen as providing indirect support for the equilibrium-based "structural" model, in that the more accurate model-implied predictive relations systematically fall within the wider standard error bands associated with the unrestricted regressions. This, of course, would not necessarily be the case if the assumptions underlying the "structural" model were violated.

2.4.4 Further Discussion and Interpretation

The contrast between the long-run predictability inherent in the dividend-price ratio, and the variance variables ability to predict both return and cash flow over shorter within-year horizons is intimately related to our equilibrium-based long-run risk model, and the way in which the fundamental risk factors affect the state variables.

In particular, while the dividend-price ratio dp_t loads on the long-run risk factor x_t and both of the volatility factors σ_t^2 and q_t, the expected variation ERV_t depends only on the two volatility factors σ_t^2 and q_t, and the variance risk premium VRP_t is exclusively determined by the volatility-of-volatility factor q_t. Consistent with earlier less formal model calibrations reported in the literature, our GMM-based estimates imply that the long-run risk factor is highly persistent with AR(1) coefficient equal to $\rho_x = 0.988$, while the consumption volatility factor is moderately persistent with AR(1) coefficient equal to $\rho_\sigma = 0.64$, and the consumption volatility-of-volatility factor is quickly mean-revering with AR(1) coefficient equal to $\rho_q = 0.46$.

In light of these estimates for the underlying systematic risk factors, it is therefore not surprising that the "structural" model implied return predictability regressions based on VRP_t, which depends solely on q_t, result in the most significant coefficients over relatively short 1-6 months horizon. Meanwhile, the regressions based on dp_t, which loads heavily on x_t, should show the greatest explanatory power over longer multi-year horizons, which, of course is difficult to detect statistically with the limited time span of data analyzed here. Also, whereas the variance risk premium is most significant over horizons less than 6 months, the expected variation ERV_t displays the most significant predictability over 6-12 months horizons, as the more persistent σ_t^2 process "shifts" the predictable forward.

The documented differences in the degree of cash flow predictability are most easily understood in terms of the correlations among the "structural" shocks. From the model estimates the cash flow shock is more strongly negatively correlated with the contemporaneous variance shock ($s_{d,\sigma} \propto \tilde{S}_{3,1} = -0.36$), than it is with the uncertainty shock ($s_{d,q} \propto \tilde{S}_{3,2} = -0.09$) or the long-run risk shock ($s_{d,x} \propto -\tilde{S}_{3,4} = -0.15$). Since the expected variation loads more heavily on σ_t^2 than q_t, while the dividend-price ratio and the variance risk premium are mostly determined by x_t and q_t, respectively, ERV_t will be more strongly negatively related to Δd_t than either dp_t or VRP_t. Because of the negative autocorrelation in Δd_t ($\rho_d = -0.23 < 0$), this in turn translates into the strongest positive short-run cash flow predictability results for the ERV_t predictor variable implied by the "structural" VAR.

2.5 Conclusion

We examine the joint predictability of return and dividend growth rates within a present value framework, explicitly imposing the economic equilibrium-based constraints from a long-run risk model with time-varying consumption volatility and volatility-of-volatility risk. The model clearly differentiate the long-run predictabil-

ity channels associated with the dividend-price ratio from the economic mechanisms responsible for the short-run predictability inherent in the variance risk premium and the expected return variation.

Consistent with Bansal and Yaron (2004a), our GMM-based estimates of the "structural" factor GARCH model point to a highly persistent latent long-run risk factor. Our estimates also corroborate the calibrations in Bollerslev, Tauchen, and Zhou (2009), and the notion that consumption volatility is more persistent than consumption volatility-of-volatility. In addition, the "structural" shocks identified within the model reveal that cash flow respond negatively to contemporaneous long-run growth shocks, while consumption volatility decreases with shocks to the long-run growth factor, and volatility uncertainty increases with long-run growth shocks. A new "leverage effect" whereby shocks to consumption volatility is negatively related to volatility-of-volatility also emerges from our "structural" estimation.

By allowing for much sharper and accurate inference than the procedures traditionally employed in the literature, the VAR implied by the "structural" model also provides striking new evidence on the return and cash flow predictability inherent in the data. Specifically, we find that the variance risk premium, and to a lesser extend the expected return variation, significantly predicts short-run within-year returns. On the other hand, the expected return variation, and to a lessor extend the variance risk premium, strongly predicts short-run within-year dividend growth rates. This latter finding stands in sharp contrast to the view expressed by a number of studies in the literature that cash flows are largely unpredictable.

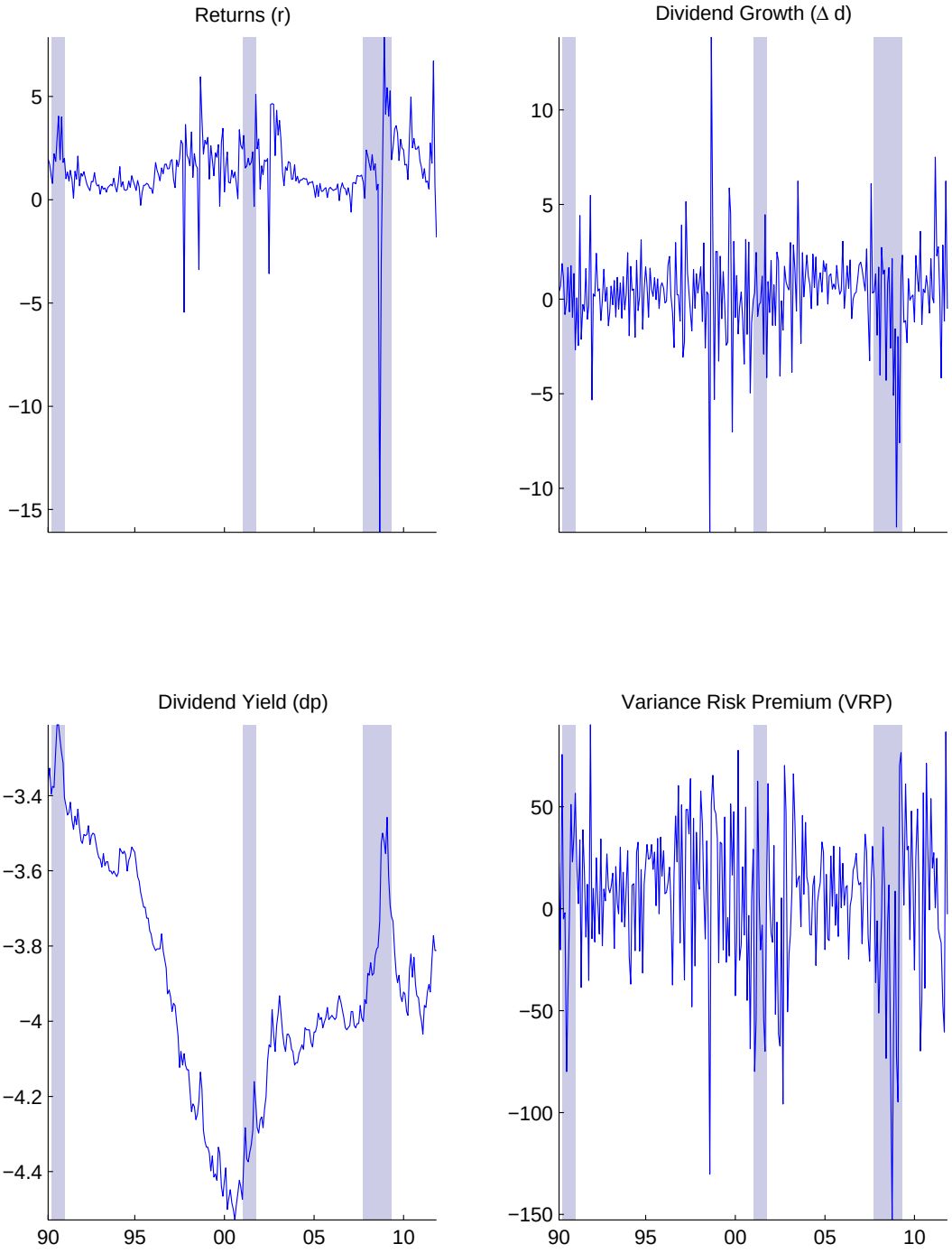


FIGURE 2.1: Returns and Dividends

The figure shows the monthly S&P500 returns (upper left panel), the log dividend growth rate (upper right panel), the log dividend-price ratio (lower left panel), and the variance risk premium (lower right panel). The returns, dividend growth, and dividend-price ratio are in annualized percentage form. The variance risk premium is in monthly percentage square form. The sample period extends from February 1990 to November 2011. The shaded areas indicate NBER dated recessions.

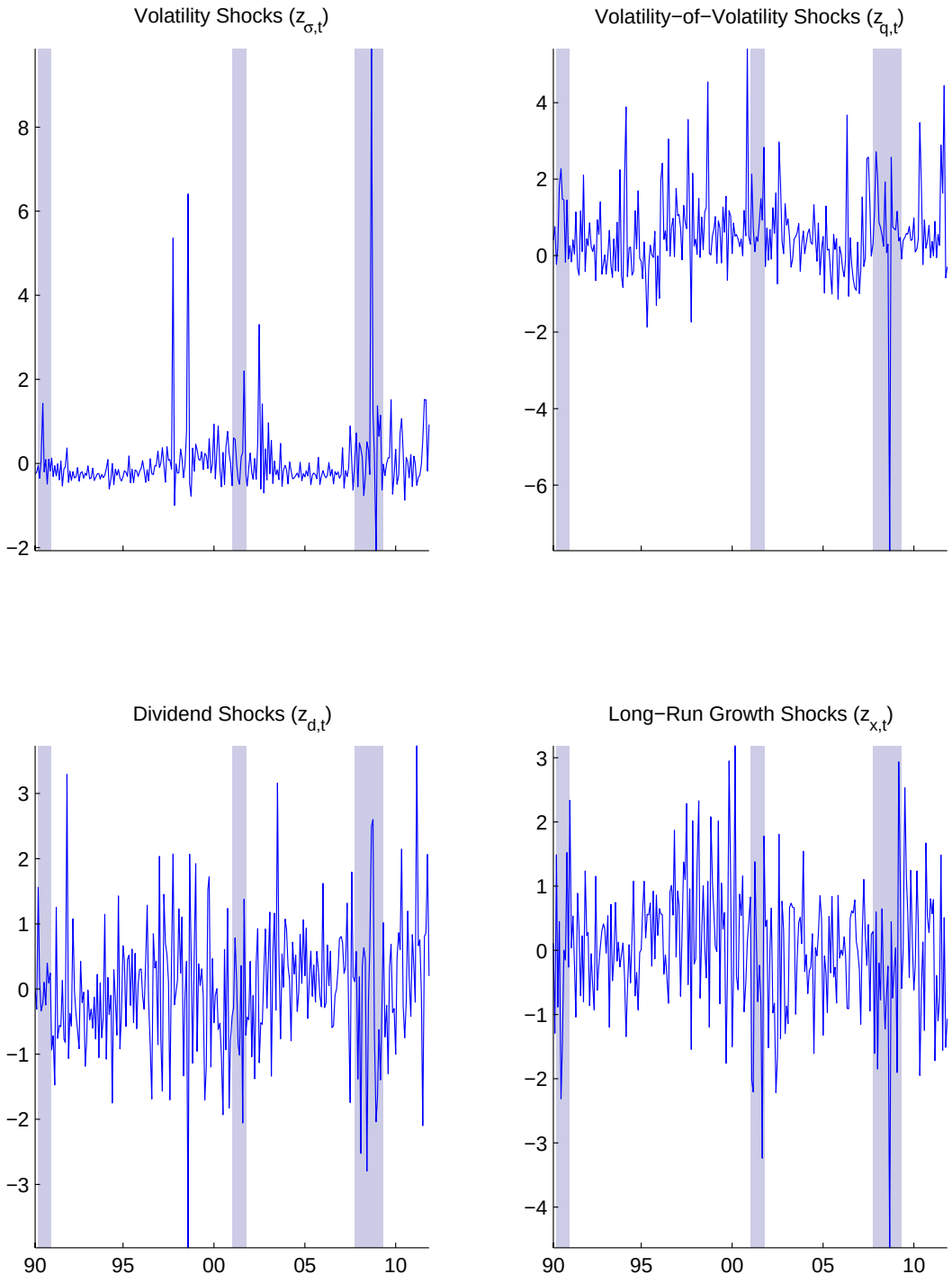


FIGURE 2.2: Model Implied Structural Shocks

The figure plots the estimated "structural" shocks z_t from the factor GARCH model discussed in the main text. The sample period extends from February 1990 to November 2011. The shaded areas indicate NBER dated recessions.

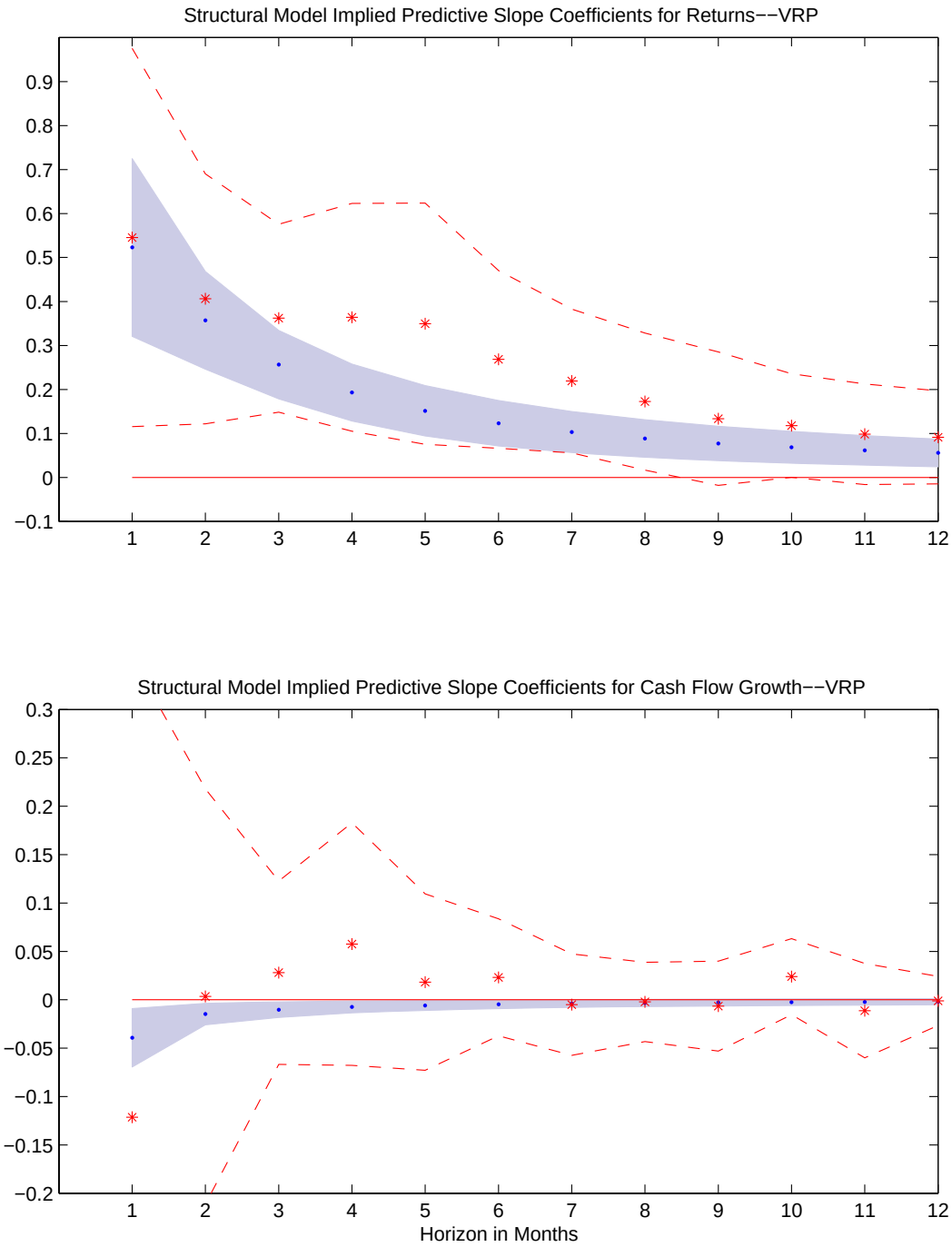


FIGURE 2.3: Predictive Regressions based on the Variance Risk Premium

The figure shows the "structural" factor GARCH model implied slope coefficients (dots) for 1-12 months return predictability regressions (upper panel) and cash flow predictability regressions (lower panel) using the variance risk premium as a predictor variable, along with 95% confidence intervals (shaded areas). The figure also shows the estimated slope coefficients from simple univariate predictability regressions using the variance risk premium as a predictor variable (stars), along with their 95% confidence intervals (dashed lines). All of the estimates are based on monthly data from February 1990 to November 2011.

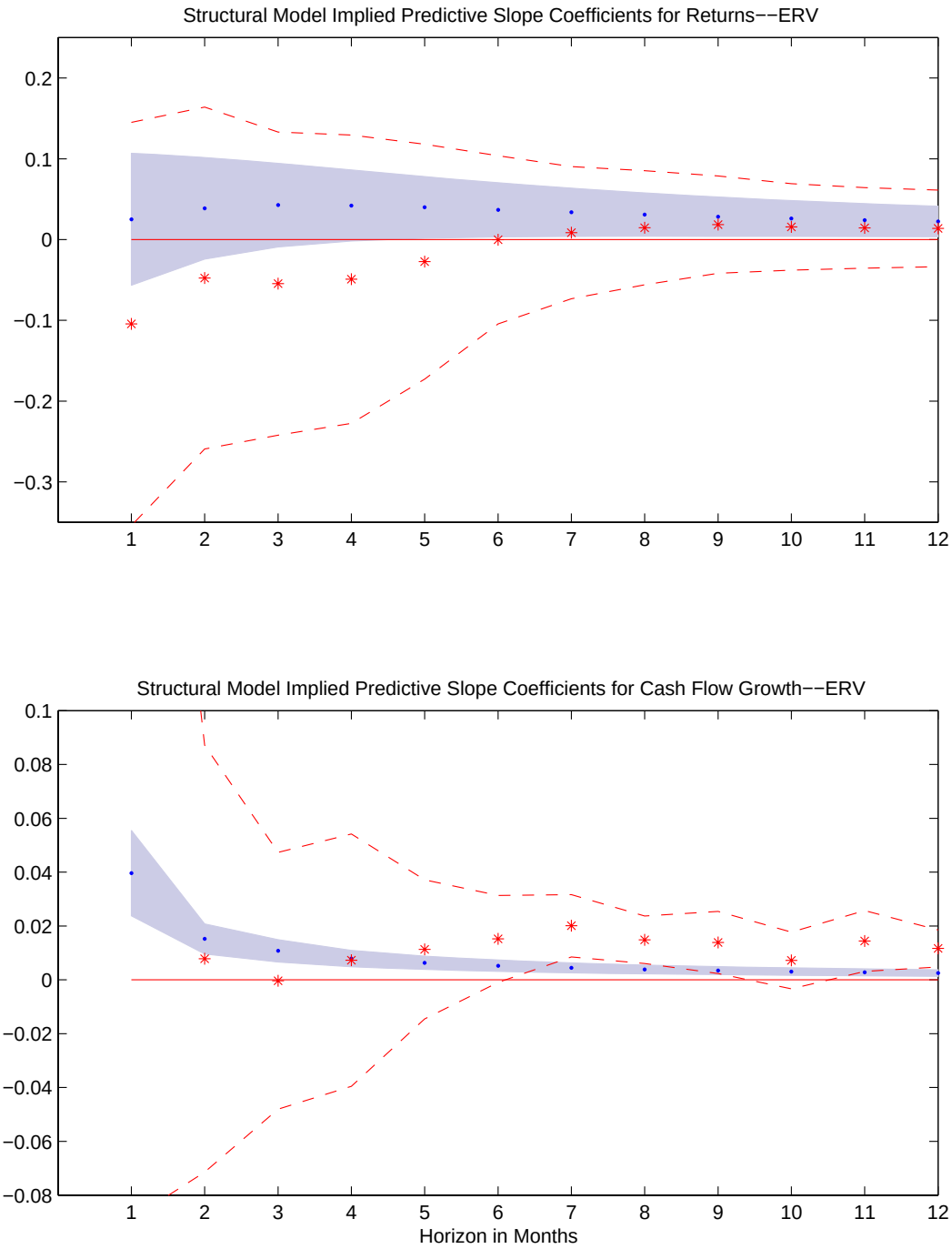


FIGURE 2.4: Predictive Regressions based on the Expected Variation

The figure shows the "structural" factor GARCH model implied slope coefficients (dots) for 1-12 months return predictability regressions (upper panel) and cash flow predictability regressions (lower panel) using the expected variation as a predictor variable, along with 95% confidence intervals (shaded areas). The figure also shows the estimated slope coefficients from simple univariate predictability regressions using the expected variation as a predictor variable (stars), along with their 95% confidence intervals (dashed red lines). All of the estimates are based on monthly data from February 1990 to November 2011.

Table 2.1: Summary Statistics

The table reports standard summary statistics and correlations for the S&P 500 return $r_{t,t+1}$, dividend growth rate Δd_t, dividend-price ratio dp_t, options implied variance IV_t, expected variance ERV_t, and variance risk premium VRP_t. The returns, dividend growth, and dividend-price ratio are all in annualized percentage form. All of the variance variables are in monthly percentage form. The sample period extends from February 1990 to November 2011, for a total of 262 monthly observations.

	Mean	Std	Skew	Kurt	AC1
$r_{t,t+1}$	8.19	15.33	-0.76	4.48	0.07
Δd_t	3.92	8.79	-0.46	10.02	-0.26
dp_t	-3.91	0.31	0.08	2.32	0.98
IV_t	40.30	36.47	3.23	18.07	0.81
ERV_t	28.54	36.64	4.62	30.08	0.69
VRP_t	11.75	14.93	-3.37	38.42	0.27

	Correlations					
	$r_{t,t+1}$	Δd_t	dp_t	IV_t	ERV_t	VRP_t
$r_{t,t+1}$	1.00	0.34	-0.03	-0.42	-0.48	0.15
Δd_t		1.00	-0.02	-0.25	-0.25	-0.01
dp_t			1.00	-0.05	-0.02	-0.07
IV_t				1.00	0.92	0.19
ERV_t					1.00	-0.21
VRP_t						1.00

Table 2.2: "Structural" Factor GARCH Estimation

The table reports the GMM estimation result for the conditional variance parameters for the "structural" factor GARCH model discussed in the main text. The column labeled ϖ_u gives the unconditional variance of the reduced form shocks u_t. Υ and Γ denote the ARCH and GARCH parameters, respectively, for the "structural" shocks $\tilde{\epsilon}_t$. The estimates are based on monthly data from February 1990 to November 2011, for a total of 262 observations.

$\tilde{\epsilon}_t$	ϖ_u		Γ		Υ	
ERV_t	0.0011	(0.0002)	0.189		0.273	(0.075)
VRP_t	0.0003	(0.0000)	0.758	(0.080)	0.239	(0.077)
Δd_t	0.0006	(0.0001)	0		0.524	(0.100)
d_{t+1}/p_t	0.0016	(0.0002)	0.299		0.163	(0.082)

Table 2.3: Predictive Regressions based on the Dividend-Price Ratio

The table reports the slope coefficients in the return and cash flow predictability regressions,

$$\frac{1}{h}\sum_{i=1}^{h} r_{t,t+i} = \alpha_{r,dp} + \beta_{r,dp}(h) \cdot dp_t + \zeta_{t,t+h}$$
$$\frac{1}{h}\sum_{i=1}^{h} \Delta d_{t,t+i} = \alpha_{\Delta d,dp} + \beta_{\Delta d,dp}(h) \cdot dp_t + \zeta_{t,t+h}$$

implied by the parameter estimates for the "structural" factor GARCH model discussed in the main text, with asymptotic standard errors in parentheses. The table also reports the slope coefficients implied by a two-variable reduced form homoskedastic VAR for the dividend growth rate and the dividend-price ratio, as in Cochrane (2008), along with the results from simple univariate predictive regressions. The time horizon h runs from one to ten years in the first two panels, and from one to twelve months in the bottom three panels. All of the results are based on monthly data from February 1990 to November 2011.

Years	1	2	3	4	5	6	7	8	9	10
					Structural Model Implied					
$\beta_{r,dp}(h)$	0.0126	0.0114	0.0106	0.0098	0.0092	0.0086	0.0081	0.0077	0.0072	0.0068
	(0.0099)	(0.0085)	(0.0074)	(0.0065)	(0.0057)	(0.0050)	(0.0045)	(0.0039)	(0.0035)	(0.0031)
$\beta_{\Delta d,dp}(h)$	-0.0013	-0.0012	-0.0011	-0.0011	-0.0010	-0.0009	-0.0009	-0.0008	-0.0008	-0.0007
	(0.0028)	(0.0027)	(0.0026)	(0.0024)	(0.0023)	(0.0022)	(0.0021)	(0.0020)	(0.0019)	(0.0018)

Months	1	2	3	4	5	6	9	12
				Structural Model Implied				
$\beta_{r,dp}(h)$	0.0166	0.0157	0.0150	0.0145	0.0141	0.0138	0.0131	0.0126
	(0.0135)	(0.0127)	(0.0122)	(0.0117)	(0.0114)	(0.0111)	(0.0104)	(0.0099)
$\beta_{\Delta d,dp}(h)$	-0.0022	-0.0017	-0.0016	-0.0015	-0.0015	-0.0014	-0.0014	-0.0013
	(0.0023)	(0.0027)	(0.0028)	(0.0029)	(0.0029)	(0.0029)	(0.0029)	(0.0028)
					Univariate Regression			
$\beta_{r,dp}(h)$	0.0112	0.0119	0.0121	0.0123	0.0129	0.0135	0.0153	0.0161
	(0.0089)	(0.0083)	(0.0078)	(0.0076)	(0.0074)	(0.0072)	(0.0069)	(0.0069)
$\beta_{\Delta d,dp}(h)$	-0.0029	-0.0011	-0.0004	-0.0001	-0.0000	0.0001	0.0002	0.0001
	(0.0030)	(0.0015)	(0.0012)	(0.0008)	(0.0007)	(0.0006)	(0.0005)	(0.0004)

Table 2.4: Predictive Regressions based on the Variance Risk Premium

The table reports the slope coefficients in the return and cash flow predictability regressions,

$$\frac{1}{h}\sum_{i=1}^{h} r_{t,t+i} = \alpha_{r,VRP} + \beta_{r,VRP}(h) \cdot VRP_t + \zeta_{t,t+h}$$
$$\frac{1}{h}\sum_{i=1}^{h} \Delta d_{t,t+i} = \alpha_{\Delta d,VRP} + \beta_{\Delta d,VRP}(h) \cdot VRP_t + \zeta_{t,t+h}$$

implied by the parameter estimates for the "structural" factor GARCH model discussed in the main text, with asymptotic standard errors in parentheses. The table also reports the slope coefficients from simple univariate predictive regressions. The time horizon h runs from one to twelve months. All of the results are based on monthly data from February 1990 to November 2011.

Months	1	2	3	4	5	6	9	12
				Structural Model Implied				
$\beta_{r,VRP}(h)$	0.5228	0.3571	0.2564	0.1929	0.1514	0.1231	0.0772	0.0557
	(0.1031)	(0.0566)	(0.0396)	(0.0330)	(0.0292)	(0.0263)	(0.0200)	(0.0161)
$\beta_{\Delta d,VRP}(h)$	-0.0393	-0.0147	-0.0103	-0.0074	-0.0058	-0.0047	-0.0029	-0.0020
	(0.0154)	(0.0057)	(0.0041)	(0.0031)	(0.0026)	(0.0023)	(0.0018)	(0.0016)
				Univariate Regression				
$\beta_{r,VRP}(h)$	0.5454	0.4060	0.3620	0.3640	0.3494	0.2683	0.1335	0.0911
	(0.2194)	(0.1450)	(0.1090)	(0.1321)	(0.1400)	(0.1031)	(0.0773)	(0.0540)
$\beta_{\Delta d,VRP}(h)$	-0.1215	0.0035	0.0280	0.0576	0.0183	0.0232	-0.0066	-0.0012
	(0.2420)	(0.1094)	(0.0483)	(0.0640)	(0.0465)	(0.0309)	(0.0237)	(0.0130)

Table 2.5: Predictive Regressions based on the Expected Variation

The table reports the slope coefficients in the return and cash flow predictability regressions,

$$\frac{1}{h}\sum_{i=1}^{h} r_{t,t+i} = \alpha_{r,ERV} + \beta_{r,ERV}(h) \cdot ERV_t + \zeta_{t,t+h}$$
$$\frac{1}{h}\sum_{i=1}^{h} \Delta d_{t,t+i} = \alpha_{\Delta d,ERV} + \beta_{\Delta d,ERV}(h) \cdot ERV_t + \zeta_{t,t+h}$$

implied by the parameter estimates for the "structural" factor GARCH model discussed in the main text, with asymptotic standard errors in parentheses. The table also reports the slope coefficients from simple univariate predictive regressions. The time horizon h runs from one to twelve months. All of the results are based on monthly data from February 1990 to November 2011.

Months	1	2	3	4	5	6	9	12
				Structural Model Implied				
$\beta_{r,ERV}(h)$	0.0251	0.0387	0.0426	0.0422	0.0398	0.0368	0.0283	0.0223
	(0.0417)	(0.0321)	(0.0264)	(0.0225)	(0.0195)	(0.0171)	(0.0124)	(0.0096)
$\beta_{\Delta d,ERV}(h)$	0.0396	0.0152	0.0107	0.0079	0.0063	0.0052	0.0034	0.0025
	(0.0081)	(0.0029)	(0.0021)	(0.0016)	(0.0013)	(0.0011)	(0.0008)	(0.0006)
				Univariate Regression				
$\beta_{r,ERV}(h)$	-0.1044	-0.0477	-0.0546	-0.0491	-0.0274	-0.0003	0.0185	0.0137
	(0.1273)	(0.1080)	(0.0957)	(0.0911)	(0.0742)	(0.0531)	(0.0307)	(0.0242)
$\beta_{\Delta d,ERV}(h)$	0.1173	0.0079	-0.0004	0.0073	0.0113	0.0152	0.0139	0.0117
	(0.1051)	(0.0404)	(0.0243)	(0.0239)	(0.0132)	(0.0082)	(0.0059)	(0.0035)

3

Tail Risk and Equity Risk Premia

3.1 Introduction

The equity risk premium—the expected return of the equity market in excess of the risk-free rate—is intimately linked with the equity's risk exposure: intuitively, the higher the equity's riskiness, the higher the risk premium should be to compensate. To study the temporal variation in this risk premium, researchers have recently decomposed it into two components: jump tail risk and diffusive risk. This separation reveals investors' different perceptions of the likelihood of infrequent large jumps and continuous diffusive movements. However, it remains an open question as to how important the time-varying jump behavior in asset prices is for overall risk compensation. This paper attempts to shed light on this issue by proposing a novel semi-parametric estimation method for the time series of both the jump and the volatility risk premia.

The contribution of this paper is threefold. First, I present a novel dynamic model for rare jumps which relaxes several restrictions used in previous studies. The model features a self-exciting process for the jump intensity (also known as the jump arrival

rate) and allows the jump shape to be asymmetric and time-varying. This generalizes a standard compound Poisson process specification, which implies independent and identically distributed (iid) increments of jumps as well as strong restrictions for the dynamics of the jump intensity.[1] A self-exciting jump intensity allows for jumps that are not only path-dependent (in that jumps are clustering over time), but also allows past jumps to influence the arrival rate of the future jumps.[2] The distribution of the jump size is commonly assumed to be time-invariant and typically Gaussian. I also relax this stringent assumption and allow for Frechet-type distributions, thereby nesting a large class of (possibly asymmetric) distributions.[3]

Second, I derive closed-form solutions for the jump and volatility components of the equity risk premium in a stylized intertemporal capital asset pricing model (ICAPM). In contrast to a recent work by Campbell et al. (2013), my model is cast in continuous-time with more empirically realistic jumps. The endogenously determined jump component of the equity risk premium entails a multi-factor structure that is directly related to the jumps in total wealth. As usual, this premium can also be conveniently expressed as the difference in ex-ante expectation of jump tails between physical ($\mathbb{P}$) and risk neutral ($\mathbb{Q}$) measures.[4] Under each measure, both the time-

[1] For example, Maheu et al. (2013) assume that the jump intensity follows an auto-regressive model and exposed to independent diffusive shocks.

[2] One of Eraker's earlier work Eraker (2004) sheds some light on the self-exciting jump intensity model by assuming the volatility co-jumps with the asset price, and the jump intensity inherits the same feature from a linear relationship with the volatility. A working paper by Aït-Sahalia et al. (2013) uses the mutually exciting processes in the international equity markets. But in their model, the volatility is independent of the jump intensity. In contrast to these studies, the jump intensity here has a flexible dependence structure with the volatility.

[3] Because of this setting, the model no longer imposes the same decay rate for both positive and negative jumps, potentially manifesting in much richer dynamic properties. The jump shape, characterized by the decay rate, is the key determinant of the jump size distribution. The higher the decay rate, the lower the probability implied for a jump of a given size.

[4] There is another strand of literature on equity risk premium, which does not rely on any specified pricing kernel. The advantage of this approach is that it is potentially model-free. However, because of this, the mechanism of the overall risk measurement and especially the role of the collective investor's preference is unclear.

varying jump shape and the self-exciting jump intensity carry important nonlinear effects on the jump risk premium. On a theoretical level, the framework implies the existence of a non-negligible stochastic shape premium and stochastic intensity premium. The former is introduced by the dynamic response of the total wealth portfolio to the aggregate market portfolio, and the latter comes from the path-dependent and self-exciting jump intensity. In addition to a succinct expression for the equity risk premium, the model implies a direct relationship between $\mathbb{P}$ and $\mathbb{Q}$-measures which is exploited for estimation.

The estimation method is model-free under the risk neutral measure and semi-parametric under the physical measure.[5] Intuitively, short-maturity and deep out-of-money (OTM) options are mostly affected by jumps, allowing us to separately identify the jump risk from the diffusive risk. Based on these option panels, the jump shape parameter is uniquely non-parametrically identified by measuring the slope of option prices versus their associated moneyness; the jump intensity can then be backed out using knowledge of the jump shape parameter.[6] In contrast, under the physical measure, a similar technique is infeasible due to so-called "peso problems."[7] To overcome this difficulty, I exploit the model-implied relationship between the two probability measures. In particular, the symmetric dynamic response of total wealth return identifies the shape premium through the difference between deep OTM puts and OTM calls (after some adjustment). Based on this along with high-frequency intra-day index prices, I estimate the intensity premium without any dynamic restriction.

In the third place, a number of results emerge from the estimated time-series of

[5] The model-free approach for both measures can guard against potential mis-specification of the model designs, e.g. Bollerslev and Todorov (011b) and Du and Kapadia (2012).

[6] For similar studies, see Carr and Wu (2003), Bollerslev and Todorov (011b), Bollerslev and Todorov (2013).

[7] Even with high-frequency intra-day data, it is very unlikely to observe more than hundreds of "medium-sized" jumps in the entire sample period, not alone for more extremely large jumps.

equity risk premia. Firstly, the jump part of the equity risk premium (ERPJ) has a mean of 6.75% on an annual basis, accounting for 93% of the total risk premium. This number is much higher than earlier estimates by Eraker (2004) and Broadie et al. (2007), but is comparable to the jump risk premium reported in Bollerslev and Todorov (011b). This larger compensation for rare events is mainly induced by the self-excitation of the jump intensity. Secondly, the serial dependence in different parts of the equity risk premium is a natural channel for explaining their strong return predictability. The deeper-tail of jumps and volatility parts together can explain 6.53% of the total variation of three-month market returns. From one to six months, the associated R^2s stay well above 2.79%. This strong forecastability reflects the importance of considering the special structure of jumps separately from volatility. Based on the estimated beta, the resulting jump and volatility components of the portfolio's equity risk premium deliver strong predictive power for that portfolio returns in short horizons of one-six months. Specifically, the "Winners-Losers" portfolio (WML) is well explained by the deeper-tail of jump and the volatility parts of the WML's equity risk premium with R^2 up to 34.88% at a half-year horizon.

There is a large related literature pertinent to the models used and modified in this paper. The closed-form model solution is similar in form to consumption based models such as "rare-disaster" or "long-run risk." Wachter (2013) uses a jump process to model the disaster events for consumption and she shows that the equity risk premium depends on the time-varying jump risk. Drechsler and Yaron (2011b) assume the consumption growth process is smooth, but the volatility of short-run consumption growth is exposed to a jump component which then implies a jump risk premium in the equity market. In contrast to these models, my extended ICAPM postulates that investors take certain types of risk in asset price as given, and then choose their consumption to satisfy the budget constraints, rather than the other way around. Since the goal of this paper is to estimate the equity risk premium,

ICAPM conveniently avoids the use of consumption data which is not measured in high enough frequency.

The estimation procedure is related to earlier studies using a variety of jump diffusion models to jointly explain options and the underlying stock price dynamics in a unified framework. These models typically rely on specific parametric assumptions: e.g. Pan (2002) and Eraker (2004) assume that the jump intensity is affine in the stochastic volatility.[8] Santa-Clara and Yan (2010) proposes a more flexible path-dependent jump intensity model to differentiate it from the volatility and argue that jump risk is more important than diffusive risk.[9] A recent paper by Li and Zinna (2013) seeks to estimate a self-exciting dynamic for the jump intensity, while at the same time allowing for volatility jumps.

In studying the time-varying jump shape, one strand of literature uses daily asset returns to show that the power-law parameter may change over time, e.g. Galbraith and Zernov (2004) applies the idea in the equity index, while Kelly (2011) relies on a large cross-section of stock returns. Another strand employs option prices to estimate the parameters of a Generalized Pareto Distribution, e.g. Hamidieh (2012), Vilkov and Xiao (2013). In contrast to the above studies, I only require short-maturity deep out-of-money option panels to estimate the jump shape under both risk neutral and physical measures. This is a benefit of having a model-implied closed-form pricing kernel.

My empirical results contribute to the studies of short-run return predictability.

[8] Most of the options literature estimates a parametric model on options data and then evaluates the fit via model implied asset prices, e.g. Bates (1996), Bates (2000), Bakshi et al. (1997). These studies strictly rely on the stochastic volatility model by Heston (1993) in which the jump intensity is either constant or affine in volatility. Broadie et al. (2007) employs both options and the index price in the estimation; their proposed model allows volatility jump but the jump intensity is constant.

[9] Maheu et al. (2013) uses the asset price only in a much longer time window; they also model the jump intensity separately in addition to a two-factor volatility structure. Both Santa-Clara and Yan (2010) and Maheu et al. (2013) successfully obtain significant equity risk premiums associated with different latent factors.

The variance risk premium (the difference between the statistical and risk-neutral expectation of the corresponding forward variation) shows strong return predictability at quarterly horizon, first documented by Bollerslev et al. (2009), further investigated by Drechsler and Yaron (2011b), Bollerslev et al. (2011), among others. Following Li and Zinna (2013), the decomposition (volatility and jump parts) of the variance risk premium significantly improves their forecasting power and the degree of this predictability has a hump shape pattern peaking at three months. However, my decomposition is based on a semi-parametric estimation procedure with only short-dated options, while Li and Zinna (2013) require a tightly specified parametric model and information about the term structure of the variance swap rates.

The rest of the paper is organized as follows: Section 2 presents the general setting for asset return process, section 3 shows an extended ICAPM with jumps, section 4 discusses the estimation strategy, section 5 describes the data and estimation results, and section 6 concludes.

3.2 Asset Return Dynamics

To study the risk premium for the aggregate market and individual equities, I set up a model for the distributional properties of asset returns. This approach is quite general and forms the foundation of the structural model I later present in section 3.

3.2.1 Rare Jump Diffusion Model

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with information flow $(\mathcal{F}_t)_{t>=0}$. On this space, I model the cumulative return on the aggregate equity market R_t as a rare jump

diffusion process satisfying,[10]

$$\frac{dR_t}{R_{t^-}} = a_t dt + \sigma_{m,t} dW_{m,t} + \int_{\mathbb{B}} (e^x - 1)\tilde{J}(dt, dx). \tag{3.1}$$

where a_t refers to the instantaneous drift, $\sigma_{m,t}$ is the stochastic volatility, and $W_{m,t}$ denotes a standard Brownian motion. Both a_t and $\sigma_{m,t}$ are locally bounded cad-lag processes and left unspecified at this stage.[11] J is a random measure for counting jumps on $[0, \infty) \times \mathbb{B}$, with a predictable jump compensator (or intensity measure) $v_t(dx)dt$ such that $\int_{\mathbb{B}} v_t(dx)dt < \infty$. $\mathbb{B}$ is defined as a subset of the real line, $\mathbb{B} = [-\infty, -b^-] \bigcup [b^+, +\infty]$, both b^+ and b^- are positive numbers.[12] Consequently, the compensated jump measure $\tilde{J}(dt, dx) = J(dt, dx) - v_t(dx)dt$ is a martingale measure and $\int_{\mathbb{B}}(e^x - 1)\tilde{J}(dt, dx)$ is a martingale process.[13]

In this model, given the information flow $\mathcal{F}_t$, the future return is exposed to two types of shocks: a continuous martingale ($\sigma_{m,t}dW_{m,t}$), and a discontinuous martingale which is the "demeaned" sum of realized large jumps. "Large" jumps refer to extremely rare events, which I take to be the top and bottom quantiles of discontinuous movements. The bounds b^- and b^+ define these threshold quantiles. In section 4, I use option implied volatility to determine numerical values for b; this method generates sufficiently large thresholds to pass any existing jump test.[14] For

[10] The cumulative return R_t is not the asset price P_t, because R_t contains both capital gains and cash flow, $\frac{dR_t}{R_{t^-}} = \frac{dP_t + D_t dt}{P_{t^-}}$ where D_t is the dividend payout at time t.

[11] In general asset pricing models, such assumptions are widely used, see e.g. Bollerslev and Todorov (011b). Depending on the properties of total wealth return and the agent's preference, a_t can takes different functional forms in an arbitrage-free world. I provide an endogenous model solution for a_t in section 3.

[12] In the empirical sections, I let both b^+ and b^- to be both time-varying and sufficiently large.

[13] Compared to an infinite activity process, a rare jump diffusion process defines jumps as rare events and these large jumps are of finite variation. Since small jumps with possibly infinite activity are excluded in the current setup, the finite variation condition for a martingale process is naturally satisfied here.

[14] In appendix C.3, based on Barndorff-Nielsen and Shephard (2004), I test for the jump existence through intra-day high-frequency prices.

a more general setup including both large and small jumps based on Poisson random measure, see Jacod and Todorov (2010).

3.2.2 Time-varying Jump Intensity Measure

I model the jump measure $J([0,t] \times \mathbb{B})$ as depending on an underlying counting process N_t which is independent of the jump size. In this underlying process, the number of jumps per unit of time follows a Poisson process, i.e. $dN_t \sim Poisson(\lambda_t dt)$. This leads to a multiplicatively separable intensity measure $v_t(dx)dt = \lambda_t f_t(x)dxdt$, where λ_t is the instantaneous intensity of jump occurrences and $f_t(x)$ is a density function for the jump size. There are two sources of randomness in the intensity measure: the arrival rate λ_t and the density function $f_t(x)$.

To match the empirical observation that jumps are rare but typically cluster in time, I model the intensity λ_t as a stationary, path-dependent, self-exciting process,

$$d\lambda_t = \kappa_\lambda(\mu_\lambda - \lambda_t)dt + \int_{\mathbb{B}} \varphi_\lambda J(dt, dx). \tag{3.2}$$

where $\varphi_\lambda > 0$ for any $x \in \mathbb{B}$. In this specification, the intensity always jumps up when the cumulative return R_t jumps, then mean reverts until the next jump.[15] A self-exciting process differs from a Poisson process by adding a source of temporal variation which is the number of jumps itself. This implies that, in contrast to other models, the discontinuous increments in both the intensity and the returns are no longer independent.[16] It is worth noting that the intensity λ_t is defined freely from the stochastic volatility $\sigma^2_{m,t}$; however, its self-exciting feature does resemble an

[15] The counting process N_t with intensity λ_t as in equation (2.2) is also called a Hawkes process, first used by Hawkes (971b) and recently adopted by Aït-Sahalia et al. (2013). The only difference between a compound Poisson process and a Hawkes process is the independent increment assumption.

[16] Alternative specifications of the jump intensity are $\lambda_t \propto \sigma^2_{m,t}$ and $\sigma^2_{m,t}$ has no jump component (see Aït-Sahalia et al. (2012)) or that λ_t is a process independent of volatility $\sigma^2_{m,t}$ (see Maheu et al. (2013)).

ARCH effect in volatility.

At the same time, to capture the possibly time-varying distribution of jump sizes, I employ the double-exponential model of Kou and Wang (2002) with parameter α_t to describe the rate of decay,

$$f_t^{\pm}(x) = \alpha_t^{\pm}\pi^{\pm}e^{\alpha_t^{\pm}|k^{*\pm}|}e^{-\alpha_t^{\pm}|x|}. \tag{3.3}$$

where $f_t^{\pm}(x)$ denotes the density function for positive or negative jumps. Here, π^+ denotes the probability of large positive jumps and $\pi^- = 1 - \pi^+$ the probability of large negative jumps. This new source of randomness in the jump distribution has also been investigated by Bollerslev and Todorov (2013); they also use a heavy tail distribution instead of Merton-type normal distribution to accommodate the complex dynamic tail.[17]

3.2.3 Co-jumps

To model portfolio returns, I explicitly allow these returns to co-jump with the aggregate market return and response to the aggregate diffusive shocks,

$$\frac{dR_{i,t}}{R_{i,t}} = a_{i,t}dt + \beta_{i,t}^{\sigma}\sigma_{m,t}dW_{m,t} + \sigma_{i,t}dW_{i,t} + \int_{\mathbb{B}}(e^{\beta_{i,t}^{J}x} - 1)\tilde{J}(dt, dx) + \int_{\mathbb{B}_i}(e^{x_i} - 1)\tilde{J}_i(dt, dx_i). \tag{3.4}$$

where $[W_{m,t}, W_{1,t}...W_{N,t}]$ denotes an $(N+1)\times 1$ vector of mutually independent standard Brownian motions, $\tilde{J}_i$ is the compensated Poisson random measure on $[0,\infty)\times\mathbb{B}_i$ with intensity $\lambda_{i,t}$ and jump size distribution $f_t(x_i)$.

By assumption, the time-variation in beta loadings β_i^{σ} and β_i^{J} comes solely from

[17] While Bollerslev and Todorov (2013) focuses on the dynamic features under the risk neutral measure $\mathbb{Q}$, this paper tries to describe shifting jump shapes under both $\mathbb{P}$ and $\mathbb{Q}$ measures. A recent study by Vilkov and Xiao (2013) assumes jump size follows a generalized Pareto distribution.

movements in the jump shape parameters $\alpha_t^{\mathbb{Q}\pm}$,

$$\beta_{i,t}^{J} = \beta_{i,0}^{J\pm} + \frac{\beta_{i,1}^{J\pm}}{\alpha_t^{\mathbb{Q}\pm}}, \qquad \beta_{i,t}^{\sigma} = \beta_{i,0}^{\sigma} + \frac{\beta_{i,1}^{\sigma}}{\alpha_t^{\mathbb{Q}-}}. \tag{3.5}$$

where $\beta_{i,0}^{J\pm}$, $\beta_{i,1}^{J\pm}$, $\beta_{i,0}^{\sigma}$ and $\beta_{i,1}^{\sigma}$ are scalars to measure the risk exposure to jump and diffusive factors. The non-zero loadings $\beta_{i,1}^{J\pm}$ and $\beta_{i,1}^{\sigma}$ imply that the co-movements with the aggregate market will change when the jump shape parameters change.

3.3 An Inter-temporal Model with Stochastic Volatility and Jumps

To better understand the risk return trade-off and to provide further insight into the estimation of equity risk premia, I extend the endowment economy intertemporal capital asset pricing model (ICAPM) of Campbell et al. (2013) to include both stochastic volatility and rare jumps. In this setting, closed-form solutions for both the pricing kernel and the equity risk premium are possible.

3.3.1 Preferences

I assume a representative agent who has a claim over a consumption stream C_t in every period. This agent has an Epstein-Zin-Weil utility function,

$$U_t = [(1 - e^{-\delta s})C_t^{\frac{1-\gamma}{\theta}} + e^{-\delta s}(E[U_{t+s}^{1-\gamma}]|\mathcal{F}_t)^{\frac{1}{\theta}}]^{\frac{\theta}{1-\gamma}}. \tag{3.6}$$

where δ is time discount rate, γ is the risk aversion, and ψ is the inter-temporal elasticity of the substitution, $\theta = \frac{\gamma-1}{\psi^{-1}-1}$.[18] The agent then maximizes his utility over the lifetime consumption choices, which results in a Stochastic Discount Factor (SDF) that is a linear function of log-consumption lnC_t and the log-return on all

[18] This utility function collapses to a power utility when the risk aversion parameter equals the inverse of the inter-temporal elasticity of the substitution $\gamma = \frac{1}{\Psi}$. The discrete-time version of this utility is widely used in the long run risk literature pioneered by Bansal and Yaron (2004b), and then extended to an continuous-time setting in Bollerslev et al. (2012a).

invested assets $lnR_{c,t}$,[19]

$$dlnM_t = -\theta\delta dt - \theta\psi^{-1}dlnC_t + (\theta - 1)dlnR_{c,t}. \tag{3.7}$$

To keep the affine nature of the pricing kernel and therefore analytical tractability, I follow Campbell and Shiller (1988b), Eraker and Shaliastovich (2008) among others by log-linearizing $lnR_{c,t}$ around $w_t = ln(P_{c,t}/C_t)$, which gives a convenient approximation for $lnR_{c,t}$,[20]

$$dlnR_{c,t} \approx \kappa_0 dt + \kappa_1 dw_t - (1 - \kappa_1)w_t dt + dlnC_t. \tag{3.8}$$

With equations (3.7) and (3.8), I can substitute out the consumption process lnC_t. Define the hedging demand $h_t = k_0 + k_1 w_t$.[21] Then equation (3.7) collapses to the following expression,

$$dlnM_t = -\gamma dlnR_{c,t} + \frac{\theta}{\Psi}\left(dh_t - [\psi\delta - \frac{\kappa_0}{\kappa_1} + \frac{1-\kappa_1}{\kappa_1}h_t]dt\right). \tag{3.9}$$

The first term in this equation represents endowment risk (a negative shock to the return on consumption $lnR_{c,t}$) as an indicator of "bad times." An asset that provides insurance against the market downturn is valuable and thus carries a lower premium. This is because investors dislike the downside risk and pay more for an asset with hedging features. The second term arises from other types of risks that matter for the pricing kernel. To further investigate this, I construct an explicit rare jump diffusion

[19] For a detailed derivation of this formula, see appendix C.1 in Bollerslev et al. (2012a). For a discrete time version for the solution, see Campbell (1993).

[20] Engsted et al. (2012) supports the accuracy of the Campbell-Shiller approximation. Here $\kappa_1 = exp(E(w_t))(1 + exp(E(w_t))^{-1}$ and $\kappa_0 = ln[1 + exp(E(w_t)] - \kappa_1 E(w_t)$.

[21] If an equity can hedge against certain state variables, its risk premium will be adjusted accordingly. This is the key element that differentiates ICAPM from CAPM and the fundamental reason for multiple factors model. For example, Campbell et al. (2013) assume the volatility is stochastic and they find that the shock to volatility is one of the state variables that describes the investment opportunity set. Thus any portfolio that is positively correlated with the volatility shock is a hedging portfolio and should earn a lower equity risk premium.

model for the return on total wealth $R_{c,t}$.

3.3.2 Return on Consumption and Volatility Dynamics

I consider a rare jump diffusion model for the endowment return on consumption stream C_t in a similar fashion as the return on the aggregate market in equation (3.1)

$$\frac{dR_{c,t}}{R_{c,t^-}} = a_{c,t}dt + \sigma_{c,t}dW_{c,t} + \int_{\mathbb{B}} (e^{\frac{\gamma_{J,t}}{\gamma}x} - 1)\tilde{J}(dt, dx) \tag{3.10}$$

$$dq_t = \kappa_q(\mu_q - q_t)dt + \varphi_q\sqrt{q_t}dW_{q,t} \tag{3.11}$$

where $\sigma_{c,t}dW_{c,t} = \sigma_c dW_{c^\perp,t} - \varphi_c\sqrt{q_t}dW_{q,t}$, $\sigma^2_{c,t} = \sigma^2_c + \varphi^2_c q_t$, as for the market return $\sigma_{m,t}dW_{m,t} = \sigma_c dW_{c^\perp,t} - \varphi_m\sqrt{q_t}dW_{q,t} + \sigma_m dW_{m^\perp,t}$, $\sigma^2_{m,t} = \sigma^2_c + \varphi^2_m q_t + \sigma^2_m$, $dW_{c^\perp,t}, dW_{m^\perp,t}$ and $dW_{q,t}$ are mutually independent shocks. The volatility component q_t follows a Heston model, or a square root process. This captures the "leverage effect" commonly documented in the literature: volatility goes up when asset price goes down. In fact, this is why volatility is positively priced as a risk factor in the economy.[22]

The last component represents the co-jumps with the market return R_t, with proportionality parameter $\frac{\gamma_{J,t}}{\gamma}$. Previous studies have assumed that $\gamma_{J,t} = \gamma$, i.e. the jump size in total wealth return is the same as the jump size in the aggregate equity market return. In the present setup, despite dependence on a common jump counting process, the total wealth return and the aggregate equity market can have jumps of differing sizes. To capture this, I allow $\gamma_{J,t}$ to be time-varying and different from γ.

The modeling assumptions of the above specifications allow me to study total wealth return and the pricing kernel in a more flexible way, while simultaneously

[22] See appendix C.1 for the price of volatility factor q_t and appendix C.4 for a small calibration study.

capturing the fact that the aggregate equity market can be more volatile. Based on this configuration, I now study the model implied equity risk premium.

3.3.3 Decomposition of the Equity Risk Premium

I use a traditional approach found in the long-run risk literature; that is, under a no arbitrage condition, I solve the equity risk premium for the aggregate market return R_t in equation (3.1),

$$\begin{aligned} a_t &= r_{f,t} + \gamma\sigma_c^2 + \varphi_m(\gamma\varphi_c + \phi_q)q_t \\ &\quad + \lambda_t \left(\int_B (e^x - 1) f_t(x)dx - \int_B (e^x - 1) e^{-\gamma_{J,t}x + \phi_{\lambda,t}} f_t(x)dx \right). \end{aligned} \tag{3.12}$$

Defining the instantaneous equity risk premium as the difference between the drift term a_t and the risk free rate $r_{f,t}$, $ERP_t = a_t - r_{f,t}$, the total equity risk premium ERP_t then consists of two components: a volatility part $ERPV_t$ and a jump part $ERPJ_t$,

$$ERP_t = ERPV_t + ERPJ_t, \tag{3.13}$$

$$ERPV_t = \gamma\sigma_c^2 + \varphi_m(\gamma\varphi_c + \phi_q)q_t, \tag{3.14}$$

$$ERPJ_t = \lambda_t \left(\int_B (e^x - 1) f_t(x)dx - \int_B (e^x - 1) e^{-\gamma_{J,t}x + \phi_{\lambda,t}} f_t(x)dx \right). \tag{3.15}$$

This decomposition naturally represents different compensation for diffusive and jump risk. If there is no jump $\lambda_t = 0$, the equity risk premium degenerates to $ERPV_t > 0$, where the temporal variation is captured by the volatility factor q_t only. However, when jumps exist $\lambda_t > 0$, $ERPJ_t$ is added to the overall equity risk premium. On average, $ERPJ_t$ should be positive and manifest the compensation for downside rare events. This separation becomes more important when the jump intensity λ_t is no longer affine in stochastic volatility q_t, reinforcing the idea of different perceptions toward distinct sources of risk.

There is an important implication for the functional form of the jump size distribution under the risk neutral measure,

$$f_t^{\mathbb{Q}\pm}(x) = \alpha_t^{\mathbb{Q}\pm}\pi_t^{\mathbb{Q}\pm}e^{\alpha_t^{\mathbb{Q}\pm}|k^{*\pm}|}e^{-\alpha_t^{\mathbb{Q}\pm}|x|}. \tag{3.16}$$

where $\alpha_t^{\mathbb{Q}+} = \alpha_t^{+} + \gamma_{J,t}$, $\alpha_t^{\mathbb{Q}-} = \alpha_t^{-} - \gamma_{J,t}$. The probability that any negative jump occurs under the $\mathbb{Q}$ measure is a function of $\gamma_{J,t}$,

$$\pi_t^{\mathbb{Q}-}(\gamma_{J,t}) = \frac{1}{1 + \frac{\pi^+}{\pi^-}e^{-\gamma_{J,t}(k_t^{*+}-k_t^{*-})}\frac{\alpha_t^{+}\alpha_t^{\mathbb{Q}-}}{\alpha_t^{\mathbb{Q}+}\alpha_t^{-}}}. \tag{3.17}$$

If the jump size is exponentially distributed under the statistical measure, then it must also be exponentially distributed under the risk neutral measure, with the shape parameter changed by $\gamma_{J,t}$.[23] I hereby define $\gamma_{J,t}$ as the shape premium, the difference between the negative jump shape under the $\mathbb{P}$ and $\mathbb{Q}$ probability measures. This (non-zero) shape premium comes from the dynamic response of total wealth portfolios to jumps in the aggregate market with size adjustment $\gamma_{J,t}/\gamma$. If total wealth return is perfectly diversified and immune to the market jumps $\gamma_{J,t} = 0$, the shape premium will disappear and the $\mathbb{P}$ and $\mathbb{Q}$-measure probabilities of negative jumps are the same $\pi_t^{\mathbb{Q}-} = \pi^-$.

The jump intensity under the $\mathbb{Q}$ measure relates to the $\mathbb{P}$- measure intensity in a different way,

$$\lambda_t^{\mathbb{Q}} = \lambda_t \times e^{\phi_{\lambda,t}}\int_B e^{-\gamma_{J,t}x}f_t(x)dx. \tag{3.18}$$

Empirical results in section 5 suggest that the second part $\int_B e^{-\gamma_{J,t}x}f_t(x)dx$ is larger than one, and not large enough to capture the difference between the two intensities under the different probability measures. Therefore, $\phi_{\lambda,t}$ which is measuring most of this wedge, i.e. $\phi_{\lambda,t} \approx log(\lambda_t^{\mathbb{Q}}) - \log(\lambda_t)$ effectively constitutes the intensity risk

[23] Drechsler and Yaron (2011b) derive a similar result in a discrete time framework.

premium. In contrast to previous equilibrium models, these two intensities no longer share the same dynamic properties, and more importantly this intensity premium is generated from the self-excitation of the jump intensity λ_t.

For any particular portfolio return i in equation (3.4), under the no arbitrage condition, the portfolio's equity risk premium $ERP_{i,t}$=$a_{i,t} - r_{f,t}$ also consists of two parts,

$$ERP_{i,t} = ERPV_{i,t} + ERPJ_{i,t}, \tag{3.19}$$

$$ERPV_{i,t} = \beta^{\sigma}_{i,t}\left(\gamma\sigma_c^2 + \varphi_m(\gamma\varphi_c + \phi_q)q_t\right), \tag{3.20}$$

$$ERPJ_{i,t} = \lambda_t\left(\int_B (e^{\beta^J_{i,t}x} - 1)f_t(x)dx - \int_B (e^{\beta^J_{i,t}x} - 1)e^{-\gamma_{J,t}x+\phi_{\lambda,t}}f_t(x)dx\right) \tag{3.21}$$

This solution is comparable with the equity risk premium for the aggregate market in equation (3.13). The portfolio's diffusive risk premium is proportional to that of the aggregate market in a traditional way $ERPV_{i,t} = \beta^{\sigma}_{i,t}ERPV_t$. However, the jump part of the equity risk premium $ERPJ_{i,t}$ is nonlinearly related with $\beta^J_{i,t}$ and market information on the risk premium. Presumably the higher β^J_i, the higher risk compensation for that particular portfolio. I leave further discussion to the empirical results in section 3.5.

3.4 Estimation Strategy

This section introduces a new semi-parametric estimation method for the jump shape and the jump intensity. These estimates are essential to construct the jump and the diffusive risk premium. Under the risk-neutral measure, I use a direct and model-free approach based on option data only; under the statistical measure, both options and high-frequency data are needed. Lastly, the resulting jump shape and the jump intensity estimates make it possible to eliminate the jump bias embedded in the VIX index (CBOE), which helps to further reveal the volatility part of the equity risk

premium.

3.4.1 Jump Shape and Jump Intensity under the Risk Neutral Measure

Most studies assume the jump shape is fixed; in contrast, I allow it to change over time. This modeling assumption makes the estimation more challenging when the jump intensity is also time-varying. However, since the jump shape itself describes distributional properties when jumps occur, which is conceptually very different from the jump intensity (jump arrival rate), the robust estimates for both of them are equally important. Together, the jump shape and the jump intensity provide richer information about rare events from different perspectives.

The idea is to estimate the jump shape first and then the jump intensity in a model-free way. One possible model-free strategy, as shown in Bollerslev and Todorov (011b), is to use short-dated deep out-of-money options. When the expiration date is near, deep out-of-money options are only hedging against extremely large jumps.

Specifically, under the risk neutral measure, let $O_{t,T}(k)$ be the option price on the aggregate equity market at time t, with maturity date T, strike price K, and log moneyness k= $\log(K/F_{t^-,T})$, where $F_{t^-,T}$ is the futures price with some undetermined future date after T.[24] Following Bollerslev and Todorov (011b) and Bollerslev and Todorov (2013), as time-to-maturity $\tau \downarrow 0$ ($\tau = T-t$), k$\uparrow +\infty$ for calls and k$\downarrow -\infty$ for puts, the ratio of the option price to the discounted futures price effectively isolates the jump risk,[25]

$$\frac{e^{r_{t,T}} O_{t,T}(k)}{\tau F_t} \xrightarrow{p} \lambda_t^Q \int_{\mathbb{B}} (e^x - e^k)^{\pm} f_t^{\mathbb{Q}}(dx). \tag{3.22}$$

[24] This is because the only information needed here is the change of the futures price at date t and date T.

[25] See Lemma 1 in Bollerslev and Todorov (2013), in their notation $\phi_t^{\pm} = \lambda_t^{\mathbb{Q}} \alpha_t^{\mathbb{Q}\pm} \pi^{\mathbb{Q}\pm} e^{\alpha_t^{\mathbb{Q}\pm} |k^{*\pm}|}$ with boundary conditions. If $\alpha_t^- > \iota$, $\alpha_t^+ > 3 + \iota$ for some $\iota > 0$, the convergence rate is τ^ι and the cumulated return R_t has a finite third moment.

In fact, for any fixed value $\alpha^{\mathbb{Q}\pm}$, when the time-to-maturity τ decreases, the convergence rate will exponentially increase, and this rate will even be exponentially amplified when the moneyness also goes deeper. As such, in this section and the empirical discussion later on, I will ignore any estimation error if the procedure only depends on the equation (3.22).

Since the jump intensity only enters into equation (3.22) in the first order format, for any pair of options with different log moneyness k_1 and k_2, the ratio of these two option prices on the same day doesn't depend on the jump intensity,

$$log\frac{O_{t,T}(k_2)}{O_{t,T}(k_1)} \xrightarrow{p} log\frac{\int_{\mathbb{B}}(e^x-e^{k_2})^{\pm}f_t^{\mathbb{Q}\pm}(x)dx}{\int_{\mathbb{B}}(e^x-e^{k_1})^{\pm}f_t^{\mathbb{Q}\pm}(x)dx}. \tag{3.23}$$

By using the exact specification of $f_t(x)$ in equation (3.3), the right-hand-side variable includes the shape parameter $\alpha_t^{\mathbb{Q}\pm}$ and the moneyness k_1 and k_2 only, which suggests the shape parameter $\alpha_t^{\mathbb{Q}\pm}$ can be consistently estimated through a sufficiently large number of short-deep-OTM option pairs. In particular, assume on any given day s ($t \leqslant s \leqslant T_0$), the errors of downward slope (log option price versus log moneyness) are independent with a median zero conditional on $\mathcal{F}_t$. These pricing errors are also independent across different days within a short period of time. Together this suggests a least absolute difference (LAD) estimate for the jump shape,

$$\widehat{\alpha_t^{\mathbb{Q}\pm}} = \underset{\alpha_t^{\mathbb{Q}\pm}\in R}{argmin} \sum_{s=t}^{T-T_0}\sum_{i=2}^{N_s^{\pm}}\Big|\frac{log\frac{O_{s,T}(k_{s,i})}{O_{s,T}(k_{s,i-1})}}{k_{s,i}-k_{s,i-1}} - (1\mp\alpha_t^{\mathbb{Q}\pm})\Big|. \tag{3.24}$$

where $N_s^{\pm}$ is the total number of calls or puts on day s, T_0 is approximately eight calendar days right before the option expires at date T,[26] and the moneyness is sufficiently deep for both calls $k^{\alpha+} < k_{t,1} < k_{t,2}... < k_{t,N_t^+}$ and puts $k^{\alpha-} > k_{t,1} > k_{t,2}... > k_{t,N_t^-}$. This LAD estimator sufficiently down-weights the "outliers" in the

[26] I do not include observations with the maturity shorter than seven days because these options contain large amounts of noise; see discussion in Bollerslev and Todorov (011b).

minimization problem.[27] In empirical practice, the pooling on a monthly basis ($\tau \leqslant$ 35 days) ensures a large number of option pairs in the estimation.[28]

In light of equation (3.22), the jump intensity under the risk-neutral measure $\lambda_t^{\mathbb{Q}}$ is then identified based on the first-step estimated shape parameter $\alpha_t^{\mathbb{Q}\pm}$ in equation (3.24). Since the model in section 3 implies the same intensity process for both positive and negative jumps, I consequently use both calls and puts to infer the embedded information on the intensity,[29]

$$\widehat{\lambda_t^{\mathbb{Q}}} = \underset{\lambda_t^{\mathbb{Q}} \in R}{argmin} \sum_{s=t}^{T-T_0} \sum_{i=1}^{N_t^-} \sum_{j=1}^{N_t^+} \Big| \log\Big[\frac{e^{r_{s,T}} CALL_{s,T}(k_{i,s})}{\tau_s F_s} \Big/ \frac{e^{(1-\alpha_t^{\mathbb{Q}+})k_{s,j} + \alpha_t^{\mathbb{Q}+} k_t^{*+}}}{\alpha_t^{\mathbb{Q}+} - 1} + \frac{e^{r_{s,T}} PUT_{s,T}(k_{i,t})}{\tau_s F_s} \Big/ \frac{e^{(1+\alpha_t^{\mathbb{Q}-})k_{s,i} - \alpha_t^{\mathbb{Q}+} k_t^{*-}}}{\alpha_t^{\mathbb{Q}+} + 1} \Big] - \log \lambda_t^{\mathbb{Q}} \Big|. \quad (3.25)$$

where k_t^{*+} and k_t^{*-} are the thresholds to define the rare jumps on the real line. Because of the symmetry assumption, the jump intensity has an even larger sample in the estimation.

Together, the estimated asymmetric jump shapes and the symmetric jump intensity provide a full characterization for jumps under the risk neutral measure. They also prove to be helpful for inferring the jump features under the physical measure.

3.4.2 Jump Shape and Jump Intensity Premiums

The proposed ICAPM in section 3 implies an explicit relation across the two probability measures. Among others, the shape premium $\gamma_{J,t}$, or the difference between the left-jump shape under $\mathbb{P}$ and $\mathbb{Q}$ probability measures, arises from the dynamic

[27] If one relaxes the independence assumption, he or she can also try to apply the kernel weights for the absolute errors.

[28] Since I eliminate the extremely short to maturity options, $T - T_0 \neq 0$, the workable number of days in one month is 15 instead of 22.

[29] A similar approach is discussed in Bollerslev and Todorov (2013) in their footnote 13. The difference is I use both puts and calls in the estimation to reflect the independence assumption on jump intensity $\lambda_t^{\mathbb{Q}}$ and jump size.

response of total wealth to jumps in the aggregate equity market. The higher shape premium is an effect of a larger response.

There are two ways to semi-parametrically estimate the jump shape premium $\gamma_{J,t}$.[30] First, we can use high-frequency prices to estimate the $\mathbb{P}$-measure decay rates, and then compare with their $\mathbb{Q}$-measure decay rates to infer the shape premium $\gamma_{J,t}$. This is only possible if we can observe enough large jumps in the actual prices within a short period of time. In the appendix C.3, I show that for the entire sample, we can only observe about 347 "medium-sized" jumps. This means the direct use of high-frequency data is infeasible for the shape premium estimation.

The second way (feasible way) is illuminated by the relation between the shape premium and the $\mathbb{Q}$-measure probability in equation (3.17). The intuition behind this approach is the following: given a fixed physical probability for negative jumps (e.g. $\pi^- = 0.5$), the corresponding risk-neutral probability is higher than π if and only if the shape premium $\gamma_{J,t}$ is positive. In other words, $\pi^{\mathbb{Q}-}$ can uniquely identify the shape premium $\gamma_{J,t}$ based on some predetermined $\mathbb{P}$- measure probability and the $\mathbb{Q}$- measure decay rates. To be more specific, by assuming $\pi^+ = \pi^- = 0.5$,[31] it implies that $\pi_t^{\mathbb{Q}-}$ in equation (3.17) is a function of shape parameter $\alpha_t^{\mathbb{Q}\pm}$ and the shape premium $\gamma_{J,t}$ only,

$$\pi_t^{\mathbb{Q}-}(\gamma_{J,t}) \quad = \frac{1}{1+e^{-\gamma_{J,t}(k_t^{*+}-k_t^{*-})\frac{\alpha_t^+\alpha_t^{\mathbb{Q}-}}{\alpha_t^{\mathbb{Q}+}\alpha_t^-}}}. \tag{3.26}$$

Since the probability for negative jumps under the risk-neutral measure $\pi^{\mathbb{Q}-}$ can be estimated by comparing the price of puts and calls, so can the shape premium

[30] In most of the studies, shape premium $\gamma_{J,t}$ is the same as the time-invariant risk aversion parameter γ. The latter is usually a calibrated number in consumption based asset pricing models to match with the average mean of the aggregate equity risk premium; or it can be estimated in a fully parameterized jump-diffusion model. However, since I am mostly interested in reducing the model assumptions in the estimation, I have no reason to make similar restrictions at this point.

[31] From year 1990 to 2011, there are 2644 jumps in total, 1110 are positive jumps (1534 are negative jumps), among which 162 are larger than 0.6% (185 are smaller than -0.6%).

$\gamma_{J,t}$,

$$\widehat{\gamma_{J,t}} = \underset{\gamma_{J,t}\in R}{argmin} \sum_{s=t}^{T-T_0} \sum_{i=1}^{N_s^-} \sum_{j=1}^{N_s^+} \Big| \log \frac{PUT_{s,T}(k_{s,i})}{CALL_{s,T}(k_{s,j})} - (1+\alpha_t^{\mathbb{Q}-})k_{s,i} + (1-\alpha_t^{\mathbb{Q}+})k_{s,j}$$

$$+ \log(\frac{\alpha_t^{\mathbb{Q}-}+1}{\alpha_t^{\mathbb{Q}+}-1}) - \log\left(\frac{\pi_s^{\mathbb{Q}-}(\gamma_{J,t}) \quad e^{-\alpha_t^{\mathbb{Q}-}k_s^{*-}}}{\left[1-\pi_s^{\mathbb{Q}-}(\gamma_{J,t})\right] \quad e^{\alpha_t^{\mathbb{Q}+}k_s^{*+}}}\right)\Big|. \quad (3.27)$$

where the pricing errors in the put-call pairs are $\mathcal{F}_t-$ independent with a median of zero. As a result, the $\mathbb{P}$-measure shape parameter estimates are $\widehat{\alpha_t^+} = \widehat{\alpha_t^{Q+}} - \widehat{\gamma_{J,t}}$ and $\widehat{\alpha_t^-} = \widehat{\alpha_t^{Q-}} + \widehat{\gamma_{J,t}}$.

This "indirect approach" successfully overcomes the difficulty of estimating the "unobservables." This approach uses a total of $N_s^- \times N_s^+$ observations on each day. In the empirical implementation discussed below, that is approximately 114×15 observations per month. A similar approach can also be applied to the case where $\gamma_{J,t} = \gamma$ by summing the right-hand-side of equation (3.27) over the full sample horizon.

The $\mathbb{P}$-measure jump intensity estimation is complicated by a similar "lack-of-data" problem. In order to make any inference about this "unobserved" jump intensity, I assume a pseudo jump intensity measure exists with a smaller cutoff choice k^{P*+} and k^{P*-}. This means that "medium" and "large" jumps are no longer separable.

Consider a noisy proxy for the median-to-large jump intensity $E_t JX$,

$$JX^{\pm}_{k_t^{P*\pm},T} \equiv \int_t^T \int_{[k^{P*\pm},\pm\infty]} (e^x - e^{k_t^{P*\pm}}) J(dt,dx), \quad (3.28)$$

$$EJX^{\pm}_{k_t^{P*\pm},T} \equiv E_t JX^{\pm}_{k_t^{P*\pm},T} \xrightarrow{p} \lambda_t \int_{[k^{P*\pm},\pm\infty]} (e^x - e^{k_t^{P*\pm}}) f_t^{\pm}(x)dx. \quad (3.29)$$

The estimated $EJX^{\pm}_{k_t^{P*\pm},T}$ comes from a HAR-VAR Kalman filter approach, and

is based on a daily observed vector $Y = [CV, JV^{\pm}, OVER^2, JX^{\pm}]$, where CV is the continuous variation, $JV^{\pm}$ is the left- and right- jump variation, $OVER^2$ is the overnight return squared.[32] Even for medium-sized jumps, on most days $JX^{\pm}_{k_t^{P*\pm},T}$ are both zero, so the variation in $EJX^{\pm}_{k_t^{P*\pm},T}$ mostly comes from the other variables in the vector $Y_{t,T}$.

Together with equation (3.22), I then estimate $\phi_{\lambda,t}$ by taking the log difference of call prices and the right-jump tail EJX_t^+, and the log difference of put prices and the left-jump tail EJX_t^-, for any day $s \in [t, T_0]$,

$$\begin{aligned}\widehat{\phi_{\lambda,t}} = \quad \underset{\phi_{\lambda,t}\in R}{argmin} \sum_{s=t}^{T_0} \sum_{i=1}^{N_s^-} \sum_{j=1}^{N_s^+} | \log[& \frac{e^{r_{s,T}} CALL_{s,T}(k_{j,s})}{\tau_s F_s \times \widehat{JTX^+_{[s,T]}}} / \frac{e^{(1-\alpha_t^{\mathbb{Q}+})(k_{s,j}-k_t^{P*+})}}{(\alpha_t^{\mathbb{Q}+}-1)/(\alpha_t^+-1)} \\ & + \frac{e^{r_{s,T}} PUT_{s,T}(k_{i,s})}{\tau_s F_s \times \widehat{JTX^-_{[s,T]}}} / \frac{e^{(1+\alpha_t^{\mathbb{Q}-})(k_{s,i}-k_t^{P*-})}}{(\alpha_t^{\mathbb{Q}+}+1)/(\alpha_t^+ +1)}] \\ & - \log \textstyle\int_{\tilde{B}} \alpha_t^{\pm} e^{\alpha^{\pm}|k^{*P\pm}| - \gamma_{J,t}x - \alpha_t^{\pm}|x|} dx - \phi_{\lambda,t}|. \quad (3.30)\end{aligned}$$

where the pricing errors are $\mathcal{F}_t-$ independent with a median of zero. Once again, the LAD estimate downweighs the outliers. As a result, the jump intensity reveals itself naturally as $\widehat{\lambda_t} = \widehat{\lambda_t^Q} / e^{\widehat{\phi_{\lambda,t}}} / \int_B e^{-\widehat{\gamma_{J,t}}x} f_t(x)dx$. In parallel to the shape premium $\gamma_{J,t}$ estimation, this approach also uses a total of $N_s^- \times N_s^+$ observations on each day.

3.4.3 Jump Part of Equity Risk Premium

The uncovered jump shapes and intensities are building blocks in the construction of the jump risk premium. As stated in the model in section 3.3, the jump risk premium (or the jump part of the equity risk premium) contains both the risk-neutral and the physical expectation for the rare events, the difference of which justifies the compensated premium. In turn, the jump part of the equity risk premium, or

[32] For details, see section 4.4, and appendix C.3, for a similar study see Bollerslev and Todorov (011b).

any of its deeper-tails defined on a subset $\mathbb{D} \in \mathbb{B}$, $\mathbb{D} = [-\infty, k_t^{**-}] \bigcup [k_t^{***+}, \infty]$, are constructed as below,

$$\begin{aligned}
\widehat{ERPJ_{\mathbb{D},t}} &= \widehat{ERPJ^{+}_{\mathbb{D},t}} + \widehat{ERPJ^{-}_{\mathbb{D},t}}, && (3.31)\\
\widehat{ERPJ^{+}_{\mathbb{D},t}} &= \widehat{ERPI_{\mathbb{D},t}} + \widehat{\lambda}_t \int_{\mathbb{D}^+} \widehat{\alpha_t^{\pm}} \pi^{\pm} e^{\widehat{\alpha_t^{\pm}}(|k^{*\pm}|-|x|)+x} dx && (3.32)\\
&- \widehat{\lambda_t^{Q}} \int_{\mathbb{D}^+} \widehat{\alpha_t^{Q\pm}} \widehat{\pi_t^{Q\pm}} e^{\widehat{\alpha_t^{Q\pm}}(|k^{*\pm}|-|x|)+x} dx,\\
\widehat{ERPJ^{-}_{\mathbb{D},t}} &= \widehat{ERPI_{\mathbb{D},t}} + \widehat{\lambda}_t \int_{\mathbb{D}^-} \widehat{\alpha_t^{\pm}} \pi^{\pm} e^{\widehat{\alpha_t^{\pm}}(|k^{*\pm}|-|x|)+x} dx && (3.33)\\
&- \widehat{\lambda_t^{Q}} \int_{\mathbb{D}^-} \widehat{\alpha_t^{Q\pm}} \widehat{\pi_t^{Q\pm}} e^{\widehat{\alpha_t^{Q\pm}}(|k^{*\pm}|-|x|)+x} dx,\\
\widehat{ERPI_{\mathbb{D},t}} &= -\widehat{\lambda}_t \int_{\mathbb{D}} \widehat{\alpha_t^{\pm}} \pi^{\pm} e^{\widehat{\alpha_t^{\pm}}(|k^{*\pm}|-|x|)} dx && (3.34)\\
&+ \widehat{\lambda_t^{Q}} \int_{\mathbb{D}} \widehat{\alpha_t^{Q\pm}} \widehat{\pi_t^{Q\pm}} e^{\widehat{\alpha_t^{Q\pm}}(|k^{*\pm}|-|x|)} dx.
\end{aligned}$$

where if $\mathbb{D} = \mathbb{B}$, $\widehat{ERPI_{\mathbb{B},t}}$ is the difference between the risk-neutral and physical jump intensities $\widehat{ERPI_{\mathbb{B},t}} = -\widehat{\lambda}_t + \widehat{\lambda_t^{\mathbb{Q}}}$.

Compared with a fully parametric model with fixed jump shapes, so that the time-variation of the jump tails only comes from the $\mathbb{P}$- measure jump intensity, I derive a more special structure for the tail risk premium by allowing shifting jump shapes and two distinct jump intensities. Investors will be more richly compensated when the jump shape premium is larger and when the jump intensity is not only path-dependent but also self-exciting.

3.4.4 Variance Risk Premium

To show the intimate link between the variance risk premium and the volatility part of the equity risk premium, define the quadratic variation of the logarithm market

return over $[t, T]$ as,

$$QV_{t,T} \equiv \int_t^T \left(\sigma_c^2 + \sigma_m^2 + \varphi_m^2 q_s\right) ds + \int_t^T \int_B x^2 J(ds, dx). \tag{3.35}$$

Moreover, let $CV_{t,T}$ denote the continuous variation and $JV_{t,T}$ the jump variation,

$$\begin{aligned} QV_{t,T} &= CV_{t,T} + JV_{t,T}, \\ CV_{t,T} &\equiv \int_t^T \left(\sigma_c^2 + \sigma_m^2 + \varphi_m^2 q_s\right) ds, \qquad JV_{t,T} \equiv \int_t^T \int_{\mathbb{B}} x^2 J(ds, dx). \end{aligned} \tag{3.36}$$

With this notation, the variance risk premium, or the difference between the $\mathbb{Q}$ and $\mathbb{P}$-measures quadratic variation $QV_{t,T}$ is simply,

$$\begin{aligned} VRP_{t,T} &= E_t^{\mathbb{Q}}[QV_{t,T}] - E_t[QV_{t,T}] \\ &= \underbrace{E_t^{\mathbb{Q}}[CV_{t,T}] - E_t[CV_{t,T}]}_{VRP_{t,T}^{cv}} + \underbrace{E_t^{\mathbb{Q}}[JV_{t,T}] - E_t[JV_{t,T}]}_{VRP_{t,T}^{J}}. \end{aligned} \tag{3.37}$$

The implied variance IV_t quantified by VIX (CBOE), contains a bias from the jump process.[33] Removing the noise from jumps delivers a clean "model-free" measure for the continuous part of the variance risk premium,

$$\begin{aligned} \widehat{E_t^{\mathbb{Q}}[CV_{t,T}]} &= IV_{t,T} - 2\widehat{\lambda_t^{\mathbb{Q}}} \times \int_{\mathbb{B}} (e^x - x - 1)\widehat{f_t^{\mathbb{Q}}(x)} dx, \\ \widehat{VRP_{t,T}^{cv}} &= \widehat{E_t^{\mathbb{Q}}[CV_{t,T}]} - \widehat{E_t[CV_{t,T}]}. \end{aligned} \tag{3.38}$$

The recovered $\widehat{VRP_{t,T}^{cv}}$ contributes to the volatility part of the equity risk premium.[34] In appendix C.1, I prove that the continuous part of the variance risk premium

[33] Implied variance is commonly used in empirical asset pricing studies, see e.g. Bollerslev et al. (2009), Du and Kapadia (2012), among others.

[34] A parametric calibration is needed besides $\widehat{VRP_{t,T}^{cv}}$. See the calibration table in appendix C.4 for details.

$VRP^{cv}_{t,T}$ is linearly related to q_t—the driving variable for the volatility part of the equity risk premium.[35] In other words, the volatility part of the equity risk premium $ERPV_t$ is affine in the continuous part of the variance risk premium $VRP^{cv}_{t,T}$.

This section discussed a strategic plan to estimate the jump risk premium and the driving variable for the diffusive risk premium. In the next section, I present estimation results and discuss the predictive regressions for short-medium horizon returns.

3.5 Estimation Results

From the extended ICAPM with stochastic volatility and jumps in section 3, the equity risk premium is naturally decomposed into the jump part and the volatility part. In this section, I apply the estimation strategy in section 4 and present the estimates for the jump shapes and the jump intensities. I also explain the construction of the risk premia for the aggregate market and the different portfolios.

3.5.1 Data Description

S&P 500 Options

The options on the S&P 500 index are obtained from OptionMetrics. The data sample runs from January 1996 to December 2011, for a total of 4027 trading days. I remove arbitrage observations for both calls and puts based on the mid quotes, moneyness and trading volume.[36]

For each month, I use short-dated and deep out-of-money options (7-45 days time-to-maturity) with the same maturity date to ensure all options contain the information only for a fixed time period. I use deep-out-of-money options to separate

[35] In this setting, the expected continuous variation $E^Q_t[CV_{t,T}]$ and $E_t[CV_{t,T}]$ are also linearly related to q_t, however, the continuous part of the variance risk premium $VRP^{cv}_{t,T}$ is a better choice since it is robust to alternative volatility models, see e.g. the two-volatility structure in Bollerslev et al. (2009).

[36] I apply the standard cleaning rule here; for details see appendix C.2.

jumps from volatility. The "deepness" is defined by the at-the-money Black-Scholes's implied volatility σ_t^{ATM} and time-to-maturity τ; deep-out-of-money calls (puts) are those with log moneyness above (below) $k^{*+} = c^{*+} \times \sigma_t^{ATM}\sqrt{\tau}, \quad c^{*+} > 0$ ($k^{*-} = -c^{*-} \times \sigma_t^{ATM}\sqrt{\tau}, \quad c^{*-} > 0$). All together, I have on average 85 "deep" calls per month with $c^{*+} = 1.75$ and 300 "deep" puts per month with $c^{*+} = 2.0$.[37]

S&P 500 High-frequency

The high-frequency intra-day data for the S&P 500 futures are from Tick Data Inc. The data sample runs from January 1990 to December 2011. These prices are recorded every five minutes, with the first observation at 8:35 (CST) and the last one at 15:00 (CST), for a total of 78 records on each trading day.

I use high-frequency prices to non-parametrically construct the continuous variation (CV), the jump variation (JV) and the jump tail index (JX). The constructed variables are then used in a HAR-VAR Kalman filter approach to provide the expected continuous variation and the expected jump tail index EJX. I also use the daily closing prices to calculate short- and medium-horizon returns for the aggregate market.

Portfolio Returns

I employ daily returns for Small and Big companies, Growth and Value companies, and High and Low momentum portfolios. The size of the firms is determined by their market equity in June of each year. The Small (Big) firms here are those in the lowest (highest) 10 percent quantiles. The Growth (Value) firms are those with book-to-market ratio in the lowest (highest) 10 percent quantiles. For momentum portfolios, the High (Low) momentum portfolio consists of past "Winners" ("Losers"), defined by the top (bottom) 10 percent of firms based on their performance in the past 2-13

[37] For different choices of the cutoff $k^{*\pm}$, see appendix C.2. For estimation of $\alpha_t^{\mathbb{Q}\pm}$, I use a higher cutoff for puts to guarantee robust estimates: $k^{\alpha+} = k^{*+}$; $k^{\alpha-} = -3.0 \times \sigma_t^{ATM}\sqrt{\tau} < k^{*-}$.

months.[38]

In Table 3.1, the first four columns summarize statistics for the S&P 500 returns and the three Fama-French portfolios. On average, "Small-Big" earns an annual rate of 2.09%, "Value-Growth" earns an annual rate of 1.63%, and "Winners-Losers" portfolio earns an annual rate of 11.61% with the highest Sharpe ratio at 0.34.

Other Financial and Macro Variables

I also use the National Financial Condition Index (NFCI) from the Federal Reserve Bank of Chicago. This is a weekly index that provides information on financial conditions in US money markets, debts and equity markets. A positive (negative) NFCI signifies a tight (loose) financial condition. The default spread (DEF) refers to the difference between Moody's BAA bonds and AAA corporate bond yields. The term spread (TERM) is the difference between the 10-years bond yield and 3-month T-bill rate. The NFCI and all data needed to calculate the default spread and the term spread are downloaded from the Federal Reserve Bank of St. Louis.

In addition, I rely on five more macro-economic indicators. The Aruoba-Diebold-Scotti Business Conditions Index (ADS) from the Federal Reserve Bank of Philadelphia, which tracks real business conditions at a weekly frequency. A higher ADS generally indicates better real business conditions. Housing Starts from the U.S. Census Bureau, which tracks the total house start units. The consumer price index (CPI), industrial production (INDPRO), and the U-Michigan consumer sentiment (UMCSENT) are all obtained from the Federal Reserve Bank of St. Louis. All these financial and macro variables are converted to monthly frequency at the options' expiration date.

[38] These portfolios are downloaded from Professor Kenneth French's website, where detailed information is available.

3.5.2 Estimation of Jump Intensity Measures

Shifting Jump Shape

On average, the risk-neutral left-jump shape is 14.93. This is much smaller than the right-jump shape, implying a smaller decay rate for negative jumps. The shape premium is on average 3.98, indicating a positive dynamic response between the total wealth return and the aggregate equity market.[39]

The estimated jump shapes are all time-varying. Figure 3.1 plots the inverse of the estimated jump shape parameters for the left $1/\alpha_t^{\mathbb{Q}-}$(and $1/\alpha_t^{-}$) in the top panel, and for the right $1/\alpha_t^{\mathbb{Q}+}$ (and $1/\alpha_t^{+}$) in the bottom panel. The stars in the top panel show that the risk-neutral left-jump distribution clearly becomes extremely fat tailed in the late 2008 financial crisis, with an unprecedented low decay rate of $1/0.249$.[40] In the same panel, the $\mathbb{P}$-measure left-jump shape shows a similar dynamic pattern, but almost always with a faster decay rate than the corresponding $\mathbb{Q}$-measure variable, which comes from a positive shape premium, $\alpha_t^{-} = \alpha_t^{\mathbb{Q}-} + \gamma_{J,t}$. This suggests that total wealth return is almost always co-jumping with the aggregate market return in the same direction.

To better understand the source of the variation in the jump shapes, I correlate them with some selected financial indicators and macro variables in Table 3.2. The inverse of the left-jump shapes ($1/\alpha_t^{\mathbb{Q}-}$ and $1/\alpha_t^{-}$) are both highly correlated with NFCI and DEF. At the same time, these shape parameters are both negatively correlated with real business condition (ADS) and consumer sentiment (UMCSEMT). This suggests that when economics conditions worsen, the aggregate equity market is expected to have a higher probability for extremely large jumps. Conversely, the left-jump shape may be seen as signaling overall financial and real economic

[39] This number is quite similar to the calibrated risk aversion parameter $\gamma = 4$ in the rare disaster literature.

[40] This also means four moments exist for the aggregate market return process.

conditions—a heavier left tail indicates a worsened state of the economy for any type of investment.

The right-jump shape also changes over time. However, at each point of time, $\alpha_t^{\mathbb{Q}+} > \alpha_t^{\mathbb{Q}-}$ and $\alpha_t^{+} > \alpha_t^{-}$, which means under both measures, the right-jump shape is always thinner, or decays faster, than the left one. Interestingly, in late 2008, the inverse of the right-jump shape parameter has an upward spike. When the equity market suffered from a significant loss in September 2008, and news spread quickly about the multiple failures of financial institutions, a strong and quick recovery was also expected to occur in the following months.[41]

Time-varying Jump Intensity

The estimated $\mathbb{Q}$-measure jump intensity has a mean of 1.00 implying one jump every year, while the $\mathbb{P}$-measure jump intensity is much smaller with a mean of 0.21 suggesting one jump every 5 years.

In parallel to the shifting jump shapes, the jump intensities under both measures are also time-varying. As shown in Figure 3.2, the $\mathbb{P}$-measure jump intensity has several peaks over the entire sample, the largest one in late 2001 caused by the 9-11 attacks. This intensity also spikes up in other periods, e.g. the 1997-1998 Asian financial crisis, the 2002 dot-com bubble, the 2008 US financial crisis and the 2010-2011 European debt crisis.

The $\mathbb{Q}$-measure intensity (stars) is persistent (first autocorrelation equals 0.42), with a subtle descending trend from year 1996 to year 2005. Table 3.2 shows this downward trend is also correlated with industrial production (-0.53) and the consumer price index (-0.49).[42] Compared with the $\mathbb{P}$-measure jump intensity, the $\mathbb{Q}$-

[41] In general, the right-jump shape is less smooth; one possible reason is the small sample problem—for each month, there are on average 85(67) calls for the entire sample from 1996-2011 (early sample 1996-2007).

[42] The reason why the jump intensity under $\mathbb{Q}$ measure has strong negative correlation with CPI and INDPRO is not the focus of this paper, but it might be interesting for future research.

measure intensity (the stars) is almost always larger, implying a stochastic intensity premium.[43]

Based on the shape estimates and the intensity estimates under both measures, I construct the equity risk premium for the aggregate market. Consistent with the idea of risk-return trade-off, the resulting left-jump part of the equity risk premium ($ERPJ_t^-$) is almost always positive, and the right-jump part ($ERPJ_t^+$) is almost always negative. I also obtain the volatility part of the equity risk premium ($ERPV_t$) based on a small calibration study detailed in appendix C.4.[44]

3.5.3 Equity Risk Premium for the Aggregate Market

In Figure 3.3, I show the time-series of the jump and volatility parts of the equity risk premium ($ERPJ = ERPJ^+ + ERPJ^-$ and $ERPV$). On average, the jump part equals 6.75%, more than ten times larger than the volatility part on an annual basis. Both parts have peaks that are aligned with economic downturns and major financial events (the unconditional correlation between $ERPJ$ and $ERPV$ is 0.29), but in general they exhibit very different trajectories. Interestingly, the peaks in the jump part appear relatively stable over time. For example, the October 1997 Asian currency crisis is associated with a 19.03% premium from jumps, the August 1998 Russian bond and LTCM crises triggers a rise of $ERPJ$ to 20.94%, and these premia are not much different from the 2007-2008 US financial crisis when $ERPJ$ attains 19.98%. By contrast, before 2007 the volatility part has only moderate peaks, ranging from 0.56% to 1.63%, but then in Nov 2008, it soars up to 6.43% . There are additional peaks in both $ERPJ$ and $ERPV$ in connection with the September 11,

[43] To further highlight this, I also calculate the 95% confidence intervals for the estimated $\mathbb{P}$-measure jump intensity. These confidence bands are tight enough to exclude the trajectory of the $\mathbb{Q}$-measure intensity. Empirical results show that the log difference between $\mathbb{Q}$ and $\mathbb{P}$-measure intensities mostly comes from the stochastic intensity premium.

[44] Section 4 and appendix C.1 explain the relation between $ERPV_t$ and the continuous part of the variance risk premium (VRP_t^{cv}).

2001 terrorist attack, the July 2002 dot-com bubble, and the 2010-2011 European debt crisis. In all these events, the jump part of the equity risk premium captures a large increased fear for the market's decline, and the magnitude of these premia are about four times higher than the volatility part. Compared to other tail risk measures with dramatic peaks in late 2008, the magnitude of my $ERPJ$ is much smaller.[45] Since the volatility rises to unprecedented high levels in October and November 2008, the contributions from both the jump and volatility parts offset each other and become more even in comparison with other more quiet periods.

Since the jump part of the equity risk premium is constructed by the sum of the left $ERPJ_t^-$ and the right $ERPJ_t^+$ sub-parts, it is instructive to discuss the difference between this left and right decomposition for some particular event. Among all the peaks in ERPJ ($=ERPJ_t^- + ERPJ_t^+$), the left part $ERPJ_t^-$ have dominating effects over the right part $ERPJ_t^+$, and together they deliver positive risk premia. However, a reverse scenario occurs in late 2008, especially in October and November, where $ERPJ_t^-$ reaches its lowest point at -2.14% and -9.35%, and $ERPJ_t^+$ reaches its highest point at 10.46% and 29.33%. This means that the moderate positive tail risk compensation is manifest in the right jumps, not the left jumps. This counter-intuitive result is generated by the flexible dynamic response of total wealth return to the aggregate equity market $\gamma_{J,t}/\gamma$. As the total capital market value drops, the aggregate equity market is no longer a good proxy for total wealth return, inducing a sharp decrease in the shape and intensity risk premiums. Furthermore, in these two months, financial investors were under extremely tight financial conditions: NFCI rose to unprecedented high numbers of 2.6500 and 2.7400.[46]

[45] For example, the jump tail index (JTIX) in Du and Kapadia (2012) has a 50-fold increase and the jump part of the equity risk premium in Bollerslev and Todorov (011b) reaches 40% on an annual basis.

[46] National financial condition index (NFCI) summarizes the overall conditions of money markets, debt and equity markets, and these extremely large positive values of the NFCI indicate all three markets are risker, with low liquidity and high leverage.

3.5.4 Equity Risk Premia for the Fama-French Portfolios

To complement the results for the aggregate market portfolio, I also study the jump (left and right), and volatility parts of the equity risk premium for three Fama-French portfolios: SMB (small minus big firms), HML (value minus growth firms) and WML (winner minus loser firms). The constructions of these premia certainly involve the estimates of the beta loadings for each portfolio.

To start the beta pricing implications, I first split the entire data sample into three groups: the right-jump period, the left-jump period and the calm period, each of which is defined as the collection of days when the S&P500 intra-day price jumps were larger than 0.6%, smaller than -0.6%, or no such jumps, respectively.[47] The rationale for this split is to select days when the market is likely to experience large rare jumps and use those selected days to separately estimate the beta coefficients for the right-jump, left-jump and volatility. Formally, for each portfolio i, I run three regressions,

$$r_{i,t}DM_t^{J^+} = \beta_{i,0}^{J^+} \; r_{m,t}DM_t^{J^+} + e_{t,J^+}, \tag{3.39}$$

$$r_{i,t}DM_t^{J^-} = \left(\beta_{i,0}^{J^-} + \frac{\beta_{i,1}^{J^-}}{\alpha_t^{Q-}}\right) \; r_{m,t}DM_t^{J^-} + e_{t,J^-}, \tag{3.40}$$

$$r_{i,t}DM_t^{\sigma} = \left(\beta_{i,0}^{\sigma} + \frac{\beta_{i,1}^{\sigma}}{\alpha_t^{Q-}}\right) \; r_{m,t}DM_t^{\sigma} + e_{t,\sigma}. \tag{3.41}$$

where $r_{i,t}$ and $r_{m,t}$ are daily log returns, $DM_{\cdot}$ are dummy variables to indicate the three different scenarios, $\beta_{i,0}^{\sigma}$, $\beta_{i,1}^{\sigma}$,$\beta_{i,0}^{J^+}$, $\beta_{i,0}^{J^-}$ and $\beta_{i,1}^{J^-}$ are the betas for the continuous shocks, the right jumps and the left jumps in equation (3.4).

Table 3.3 reports the beta estimates for the SMB, HML and WML portfolios. For each portfolio, I have 101 days to estimate $\beta_{i,0}^{J^+}$, 99 days for $\beta_{i,0}^{J^-}$ and $\beta_{i,1}^{J^-}$, and 3746

[47] The cutoff choice 0.6% is adopted from Bollerslev and Todorov (011b). It is large enough to select large jumps and small enough to identify a reasonable number of jumps in the sample.

"normal" days for $\beta^{\sigma}_{i,0}$ and $\beta^{\sigma}_{i,1}$. In line with the literature, on average, I find that SMB, HML and WML portfolios all tend to load negatively on both the left and negative market jumps. What is more interesting here is the time-varying beta estimates in the fifth row in Table 3.3. In contrast to the right-jump beta, the left-jump beta and the volatility beta both significantly change with the jump shape parameter $\frac{1}{\alpha_t^{\mathbb{Q}-}}$. This means when financial conditions become tighter and the aggregate market has a negative jump, the Small-Big and the Value-Growth portfolio jumps in the same direction, but the Winners-Losers portfolio jumps in the opposite direction. These differences in the beta estimates naturally result in more variation in the equity risk premiums for the different portfolios.

Together with the market jump intensity and the jump shape estimates, I first investigate the resulting jump parts, then the volatility parts of the portfolios' equity risk premia. Figure 3.4 plots the right (dashed line) and left (solid line) jump parts of the equity risk premium for each portfolio in separate panels. In most months, the negative beta loadings for the right market jump for all three portfolios imply positive risk premiums. Comparing across the different portfolios, all three right-jump parts ($ERPJ^{+}_{i,t}$) are highly correlated with each other, since the variation mainly comes from the counterpart to the market $ERPJ^{+}_{t}$. Conversely, the three left-jump parts ($ERPJ^{-}_{i,t}$) are very different dynamically. This may be explained by the different stochastic betas for the negative market jumps. Specifically, in October and November of 2008, this premium almost disappears for the SMB portfolio. For HML it starts from around zero and then drops down to be negative. For WML it also starts from around zero but rises back to 10.44%. These differences come from the dual impacts of the different beta loadings and the unique market conditions. For example, in November 2008, the left-jump beta for HML is -0.0477, compared to the left-jump beta for WML equal to -1.1660, but it also has a lower associated

premium implied by the negative market-wide intensity premium.

Lastly, Figure 3.5 shows the total jump and the volatility parts of the equity risk premiums for the three portfolios. The volatility parts are almost always smaller in magnitude. Interestingly, the jump parts of the portfolios' equity risk premia exhibit very different dynamic patterns. The SMB jump premium is strongly negatively correlated with market's counterpart (correlation equals -0.95), and positively related with that of HML (0.78) and WML (0.49). The HML and WML jump premia, on the other hand, correlate negatively with the market's counterpart at -0.70 and -0.56. Unlike most of the beta pricing studies, the portfolios' equity risk premia here are not linearly related to that of the market. These additional sources of variations come from the fact that: [1] the jump shapes shift through time, and the beta loadings enter into the premium in a nonlinear fashion; [2] the beta loadings for the left-jumps also change with different financial conditions.

3.5.5 Return Predictability Studies

I now turn to the study of the long-term equity risk premium. In my model, the long-term equity risk premium (from time t to $t + n\tau$) is the summation of the expected instantaneous equity risk premium, defined as $ERP_{t,n\tau} \equiv \sum_{j=1}^{n} E_t ERP_{t+j\tau}$. If $ERP_{t,n\tau}$ can be approximated by some observables X_t, then we should have the empirical regressions,

$$\sum_{j=1}^{n} r_{t+(j-1)\tau,t+j\tau} = a(n) + b^{'}(n)X_t + \epsilon_{r,t+j\tau}. \tag{3.42}$$

where $\sum_{j=1}^{n} r_{t+(j-1)\tau,t+j\tau}$ is the compounded asset return, a noisy but unbiased proxy for the long-term risk premium. If X_t can strongly predict the future compounded asset return, then X_t is a good approximation for the long-term equity risk premium.

Table 3.4 presents these predictability studies at a one-month horizon. I try

three measures of $\widehat{ERP_t}$: my proposed measure with shifting jump shapes and self-exciting intensity (top panel), a benchmark measure with constant jump shapes (middle panel) and another benchmark measure with constant jump shapes and without self-exciting jump intensity (bottom panel). For each measure, I choose three sets of X_t in the empirical regressions: [1] the jump and the volatility parts of the equity risk premium ($ERPJ$ and $ERPV$); [2] the deeper tails of the jump part of the equity risk premium ($ERPJ(VIX)$);[48] [3] the jump part of the variance risk premium (VRP^J).

Across different columns in Table 3.4, my proposed measure for X_t has uniformly superior predictive power for the one-month S&P 500 return. For example, $ERPJ$ has R^2 equal to 1.33%, compared to 1.00% for $ERPJ^*$ and 0.96% for $ERPJ^{**}$. By adding in the volatility part in a joint regression, my proposed measure maintains the strongest predictive power of a total R^2 equal to 2.79% and adjusted R^2 equal to 1.77%. If one uses a deeper tail ($ERPJ(VIX)$) together with $ERPV$, the resulting R^2 equals 2.43%, which is 1.00% higher than the first corresponding benchmark variable ($ERPJ^*(VIX)$ and $ERPV^*$) and 1.83% higher than the second benchmark variable ($ERPJ^{**}(VIX)$ and $ERPV^{**}$). Moreover, the variance risk premium $X_t = [VRP^{cv},\ VRP^J]$ of all measures have relatively weaker return forecastability. The time-variation of the jump shapes and the self-excitation of the jump intensity are both essential to the one-month equity risk premium for the aggregate market.

This comparison also holds as the compounded horizon increases. In Figure 3.6, from one to six months, the proposed $X_t = [ERPJ \ \ ERPV]$ has uniformly stronger predictive power compared to the benchmark $X_t = [ERPJ^{**} \ \ ERPV^{**}]$. The R^2 for returns has a hump shape pattern peaking at a three-month horizon, consistent with the hump first documented by Bollerslev et al. (2009) and later by Bollerslev

[48] The deeper tail $ERPJ(VIX)$ refers to a subset $\mathbb{D}$ in equation (3.31), where $\mathbb{D} = [-\infty, k_t^{**-}] \bigcup [k_t^{**+}, \infty]$, $k_t^{**-} = -2.0 \times VIX_t\sqrt{\tau}$, $k_t^{**+} = 1.75 \times VIX_t\sqrt{\tau}$, and τ= 30 calender days.

et al. (2011). However, in the sample period considered in this paper, the forecasting power of the total VRP_t stays below 1% throughout the short-median horizon. This suggests that the tail risk is indeed very different from the diffusive risk.

Table 3.5 presents prediction results for the three portfolios Small-Big (SMB Top Panel), Value-Growth (HML Middle Panel) and Winners-Losers (WML Bottom Panel). The first four columns show that the deeper tails of the jump components explain larger amount of the portfolios' one-month equity risk premiums. The volatility parts help to increase SMB return predictability, but not significantly for HML and WML returns. In particular, the SMB portfolio can be explained by $X_t = [ERPJ_{SMB}(VIX) \;\; ERPV_{SMB}]$ with $R^2 = 7.07\%$, and the WML portfolio can be explained by $X_t = [ERPJ_{WML}(VIX) \;\; ERPV_{WML}]$ with $R^2 = 4.54\%$. The last two columns show the variance risk premium has limited contribution to the portfolios' return predictability. These results suggest that the portfolios' beta loadings are time-varying and their long-term equity risk premiums are better captured by their own instantaneous equity risk premium rather than the variance risk premium of the aggregate market.

Figure 3.7 plots the R^2s in the portfolio return predictions. For all three portfolios, the predictive power of the deeper tails are stronger than the total jump components. The R^2s are almost all increasing from one to six month horizons and no hump shape is detected at the shorter horizons.

3.6 Conclusion

This paper provides a new approach for modeling and estimating the jump risk premium in the equity market. I show that investors dislike two types of jump risk: the frequency of the discontinuous movement and the probability that this movement is extremely large. By incorporating these two different channels, and further allowing the jump arrival rate to be self-exciting, the complex structure of

the jump risk premium is no longer affine in the jump arrival rate.

The corresponding semi-parametric estimation approach developed here relies on a large panel of options and high-frequency intra-day prices to uncover the associated risk compensation. Together with the diffusive risk premium, the proposed jump risk premium is capable of providing a better description for the longer-term equity risk premium. It also helps explain differences in the returns across the three Fama-French portfolios.

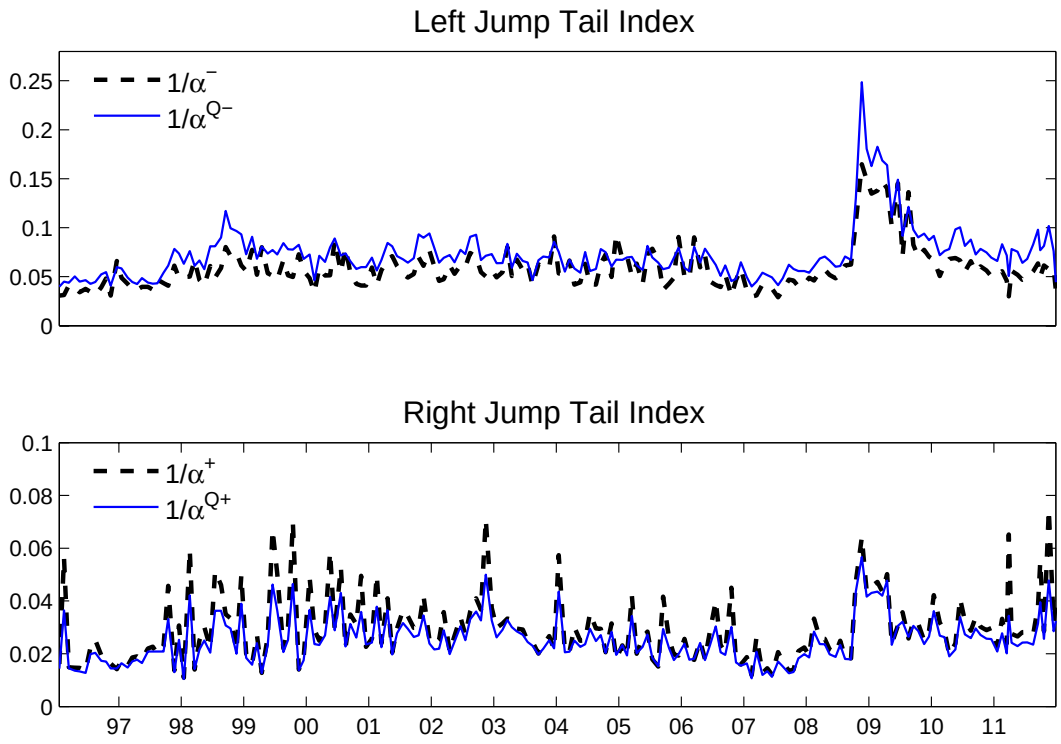


FIGURE 3.1: Jump Tail Index

This figure plots the inverse of the estimated left-jump (right-jump) shape parameters under the physical ($\mathbb{P}$) and risk-neutral ($\mathbb{Q}$) measures in the top panel (bottom panel). Section 3.4 explains the detailed estimation procedures. The sample runs from January 1996 to December 2011.

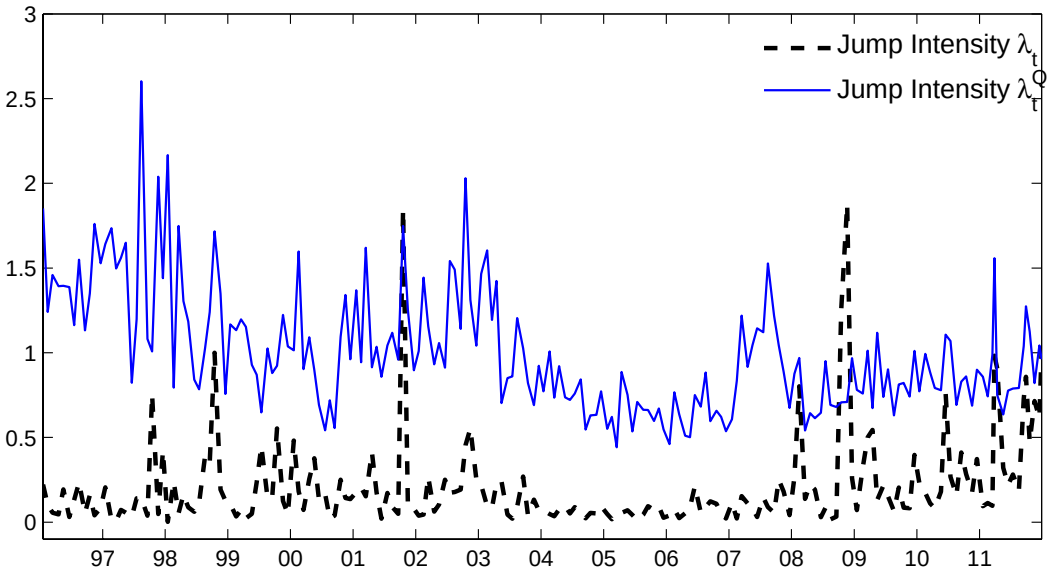


FIGURE 3.2: Jump Intensity

This figure plots the estimated jump intensity (physical measure as a dashed line and risk-neutral measure as a solid line). The shaded area represents the 2-standard error bands for the physical jump intensity. Section 3.4 explains the detailed estimation procedures. The data sample runs from January 1996 to December 2011.

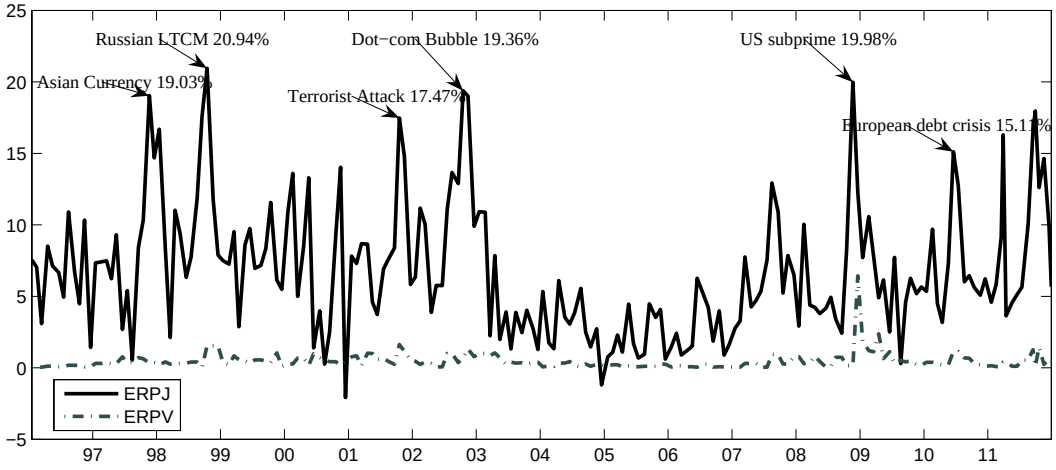


FIGURE 3.3: Decomposition of Equity Risk Premium

This figure plots the jump part of the equity risk premium $ERPJ_t$ as a solid line and the volatility part of the equity risk premium $ERPV_t$ as a dashed line. The data runs from January 1996 to December 2011.

Table 3.1: Summary Statistics

This table reports summary statistics and correlations for monthly returns of the S&P 500, SMB (Small-Big), HML (High-Low Book-to-Market) and WML (Winners-Losers), their respective jump parts of the equity risk premium, the continuous and the jump parts of the variance risk premium, and VIX (CBOE). The data sample ranges from January 1996 to December 2011. All variables are in annualized percentages.

Panel A: Summary Statistics

	SP500	SMB	HML	WML	ERPJ_{SP500}	ERPJ_{SMB}	ERPJ_{HML}	ERPJ_{WML}	VRP^{cv}	VRP^{J}	VIX
Mean	4.76	2.09	1.63	11.61	6.75	-1.72	-1.66	-0.22	2.04	1.96	6.24
Std. Dev	19.00	17.98	19.72	34.13	1.31	0.43	0.35	0.86	0.81	0.51	1.92
Skewness	-1.36	-0.49	-0.17	-0.99	0.88	-1.31	-0.25	-2.11	5.97	0.68	4.43
Kurtosis	7.51	8.24	5.81	8.32	3.56	6.55	5.48	10.66	57.81	15.14	30.47
Max	156.58	260.76	255.60	380.16	20.95	1.67	3.05	8.49	30.89	12.09	58.02
Min	-344.30	-325.44	-255.60	-606.96	-2.08	-9.59	-6.43	-17.38	0.00	-8.71	0.93
AC1	-0.04	0.17	0.07	0.04	0.54	0.41	0.38	0.65	0.31	0.38	0.70

Panel B: Correlation Matrix

	SP500	SMB	HML	WML	ERPJ_{SP500}	ERPJ_{SMB}	ERPJ_{HML}	ERPJ_{WML}	VRP^{cv}	VRP^{J}	VIX
SP500	1.00	-0.01	0.23	-0.35	-0.10	0.15	0.12	0.06	0.25	0.14	0.04
SMB		1.00	0.36	-0.06	-0.26	0.28	0.18	0.18	-0.02	-0.03	-0.18
HML			1.00	-0.44	-0.24	0.31	0.27	0.08	0.17	0.13	-0.05
WML				1.00	0.04	-0.13	-0.19	0.09	-0.36	-0.28	-0.17
ERPJ_{SP500}					1.00	-0.95	-0.70	-0.56	0.29	0.44	0.51
ERPJ_{SMB}						1.00	0.78	0.49	-0.13	-0.15	-0.43
ERPJ_{HML}							1.00	-0.16	0.21	0.04	-0.04
ERPJ_{WML}								1.00	-0.56	-0.38	-0.69
VRP^{cv}									1.00	0.61	0.75
VRP^{J}										1.00	0.41
VIX											1.00

Table 3.2: Jump Intensity Measures and Economic Indicators

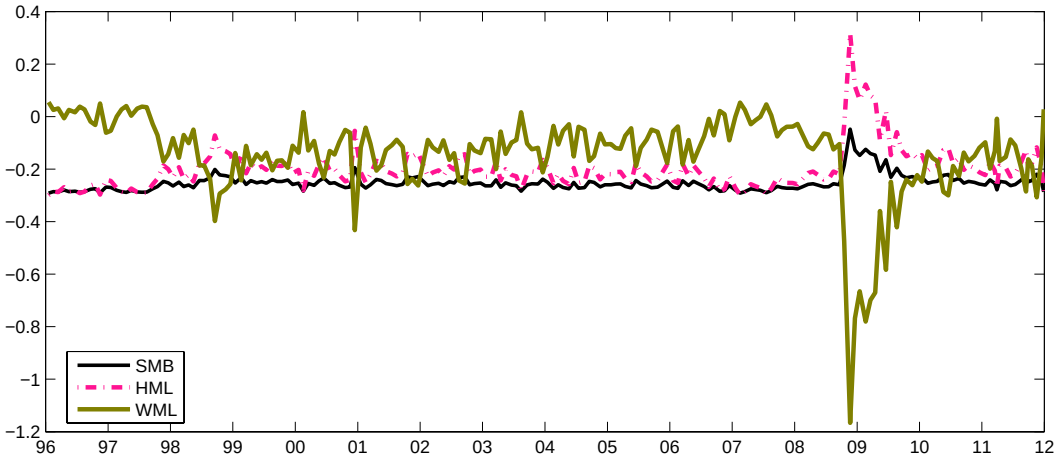
This table reports correlations for the left-jump shapes and two jump intensities with some selected financial variables: National Financial Condition Index (NFCI), the default spread (DEF), the term spread (TERM); and with some selected macro-variables: total houses start units (HouseStart), consumer price index (CPI), industrial production (INDPRO). The sample ranges from January 1996 to December 2011.

	NFCI	DEF	TERM	ADS	HouseStart	CPI	INDPRO	UMCSENT
$\frac{1}{\alpha_t^{\mathbb{Q}-}}$	0.75	0.75	0.28	-0.56	-0.39	0.27	-0.01	-0.40
$\frac{1}{\alpha_t^{-}}$	0.65	0.68	0.24	-0.52	-0.27	0.22	-0.01	-0.32
$\lambda_t^{\mathbb{Q}}$	-0.13	-0.19	-0.07	0.20	0.04	-0.49	-0.53	0.27
λ_t	0.42	0.34	0.18	-0.20	-0.35	0.22	0.04	-0.33

Table 3.3: Beta Loadings

The top panel reports the beta estimates for SMB (left panel), HML (middle panel) and WML (right panel) portfolios, for each portfolio, I estimate right-jump beta (J^+), left-jump beta (J^-) and calm period beta (V). For each type, $\beta_t = \beta_0 + \beta_1/\alpha_t^{\mathbb{Q}-}$ in equation (3.39). The bottom panel is the time-series plot for the left-jump betas, $\beta_{i,t}^{J-} = \beta_{i,0}^{J-} + \beta_{i,1}^{J-}/\alpha_t^{\mathbb{Q}-}$. Daily portfolio returns come from professor Kenneth French's website and high-frequency data for the S&P 500 futures come from TAQ dataset. The sample runs from 01Jan1996 to 31Dec2011.

	SMB			HML			WML		
	J^+	J^-	V	J^+	J^-	V	J^+	J^-	V
Constant	-0.009	0.004	0.032	-0.003	-0.001	0.017	0.009	0.000	0.040
std	(0.003)	(0.002)	(0.015)	(0.004)	(0.002)	(0.016)	(0.008)	(0.004)	(0.029)
β_0	-0.434	-0.338	-0.512	-0.135	-0.417	-0.375	-0.818	0.288	0.379
std	(0.079)	(0.099)	(0.055)	(0.098)	(0.157)	(0.083)	(0.163)	(0.234)	(0.177)
β_1		1.169	1.831		2.951	3.837		-5.849	-7.386
std		(0.694)	(0.530)		(0.987)	(0.869)		(1.669)	(2.150)
$R2$	40.494	21.766	25.256	3.196	10.097	4.467	24.723	26.412	9.271



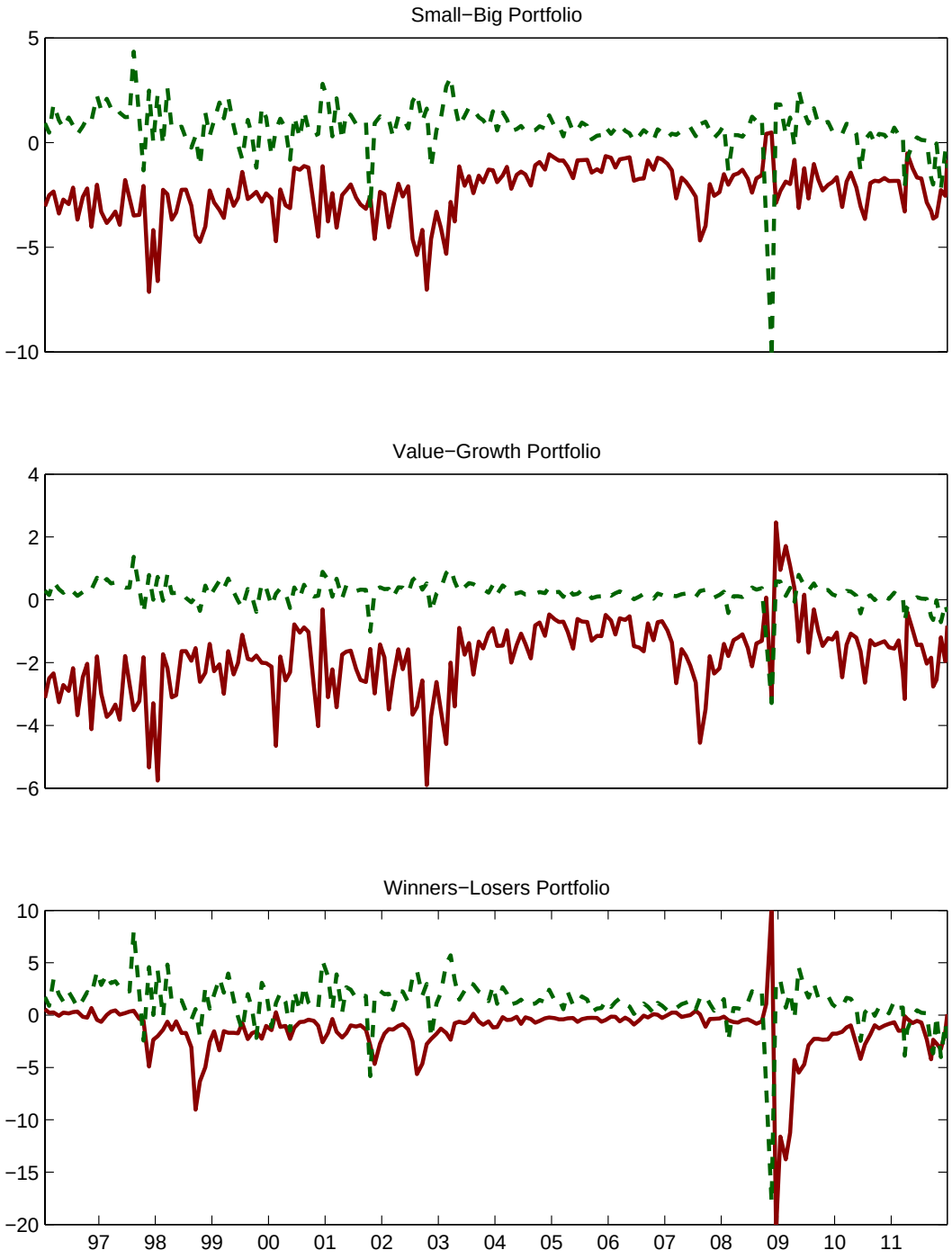


FIGURE 3.4: Decomposition of Jump parts of Portfolios' Equity Risk Premium

This figure plots right and left jump risk premiums for SMB, HML and WML. The constructions are based on tail shape parameters $\frac{1}{\alpha_t^{\mathbb{Q}\pm}}$ and $\frac{1}{\alpha_t^{\pm}}$, the intensities $\lambda_t^{\mathbb{Q}}$ and λ_t, and the beta estimates in Table 3.3. The sample runs from January 1996 to December 2011.

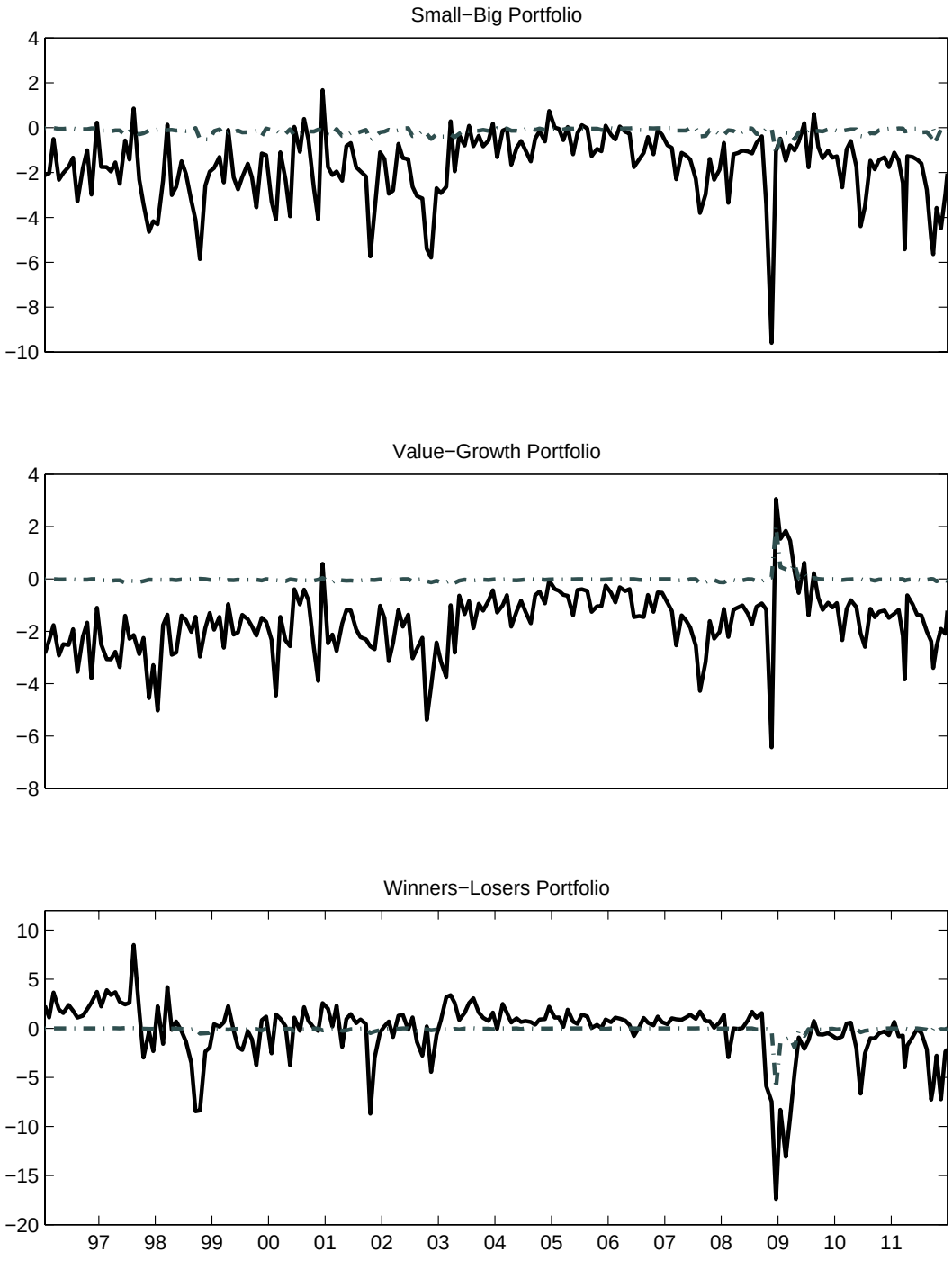


FIGURE 3.5: Decomposition of Portfolios' Equity Risk Premium

This figure plots the jump and diffusive risk premiums for SMB, HML and WML. The construction are based on tail shape parameters $\frac{1}{\alpha_t^{\mathbb{Q}\pm}}$ and $\frac{1}{\alpha_t^{\pm}}$, the intensities $\lambda_t^{\mathbb{Q}}$ and λ_t, and the beta estimates in Table 3.3. The sample runs from January 1996 to December 2011.

Table 3.4: One-Month S&P 500 Return Prediction

This table reports the compounded return predictions based on three measures (with various components) of the equity risk premium. The return process is S&P 500; ERP represents the proposed measure with the shifting jump shapes and the self-exciting jump intensity; ERP* restricts the shape to be constant, and ERP** further restricts the jump intensity to be not self-exciting. The constants and their associated standard errors are scaled by 100. The data sample ranges from January 1996 to December 2011. All measures of the equity risk premiums are in monthly units.

	S&P 500 Returns					
Constant	-0.57	-0.80	-0.37	-0.59	0.19	0.23
std	(0.77)	(0.73)	(0.74)	(0.71)	(0.54)	(0.50)
ERPJ	1.67		2.19			
std	(0.83)		(0.92)			
ERPJ(VIX)		2.86		3.41		
std		(1.28)		(1.31)		
ERPV			-2.92	-2.54		-3.68
std			(1.12)	(1.09)		(1.56)
VRP^J					1.08	4.65
std					(2.96)	(3.35)
R^2	1.33	1.29	2.79	2.43	0.08	1.67
Adjusted R^2	0.82	0.78	1.77	1.41	-0.44	0.64
Constant	0.01	-0.17	0.56	0.22	0.12	0.67
std	(0.42)	(0.53)	(0.53)	(0.58)	(0.42)	(0.53)
ERPJ*	0.46		0.66			
std	(0.23)		(0.26)			
ERPJ(VIX)*		1.21		1.51		
std		(1.25)		(1.35)		
ERPV*			-4.91	-3.75		-4.44
std			(3.55)	(3.35)		(3.55)
VRP^J*					0.95	1.29
std					(0.43)	(0.47)
R^2	1.00	0.63	2.28	1.44	0.87	1.95
Adjusted R^2	0.49	0.11	1.26	0.41	0.35	0.92
Constant	0.05	0.25	0.61	0.61	-0.11	0.39
std	(0.42)	(0.56)	(0.53)	(0.61)	(0.45)	(0.53)
ERPJ**	0.43		0.60			
std	(0.20)		(0.23)			
ERPJ(VIX)**		0.31		0.54		
std		(1.76)		(1.89)		
ERPV**			-4.84	-3.10		-5.02
std			(3.56)	(3.30)		(3.55)
VRP^J**					7.70	11.04
std					(4.28)	(4.85)
R^2	0.96	0.03	2.21	0.60	1.05	2.37
Adjusted R^2	0.45	-0.49	1.18	-0.45	0.53	1.35

Table 3.5: One-Month Portfolio Returns Prediction

This table reports the compounded returns predictions based on three measures (with various components) of the equity risk premium. The return processes are the SMB, HML and WML portfolios. ERP represents the proposed measure with the shifting jump shapes and the self-exciting jump intensity. The constants and their associated standard errors are scaled by 100. The data samples range from January 1996 to December 2011. All measures of the equity risk premium are in monthly units.

	SMB Returns					
Constant	1.14	2.36	0.67	1.91	0.18	-0.05
std	(0.50)	(0.72)	(0.51)	(0.62)	(0.48)	(0.50)
ERPJ_{SMB}	6.75		7.83			
std	(2.41)		(2.94)			
ERPJ_{SMB}(VIX)		19.11		20.46		
std		(5.74)		(6.17)		
ERPV_{SMB}			-10.61	-10.29		-10.41
std			(6.75)	(6.17)		(8.57)
VRP^J					-0.04	-2.40
std					(2.14)	(3.39)
R2	2.65	5.70	4.06	7.07	0.00	0.98
Adjusted R^2	2.14	5.21	3.06	6.10	-0.52	-0.06

	HML Returns					
Constant	0.30	1.51	0.49	2.01	0.88	0.94
std	(1.11)	(0.84)	(1.19)	(0.90)	(0.55)	(0.59)
ERPJ_{HML}	1.20		2.77			
std	(6.85)		(7.68)			
ERPJ_{HML}(VIX)		14.86		20.65		
std		(7.28)		(8.45)		
ERPV_{HML}			-5.11	-10.69		1.56
std			(5.36)	(3.64)		(3.79)
VRP^J					-4.61	-4.92
std					(2.93)	(3.32)
R^2	0.04	2.04	0.33	3.31	1.42	1.45
Adjusted R^2	-0.48	1.53	-0.71	2.30	0.90	0.41

	WML Returns					
Constant	1.09	1.80	1.29	1.80	1.91	1.33
std	(0.67)	(0.60)	(0.70)	(0.59)	(1.10)	(1.17)
ERPJ_{WML}	7.11		4.87			
std	(4.92)		(5.91)			
ERPJ_{WML}(VIX)		13.17		10.70		
std		(8.27)		(10.72)		
ERPV_{WML}			4.69	3.18		8.62
std			(5.59)	(5.42)		(3.31)
VRP^J					-5.76	0.45
std					(7.24)	(7.66)
R^2	3.19	4.29	3.74	4.54	0.74	2.80
Adjusted R^2	2.69	3.79	2.73	3.54	0.22	1.78

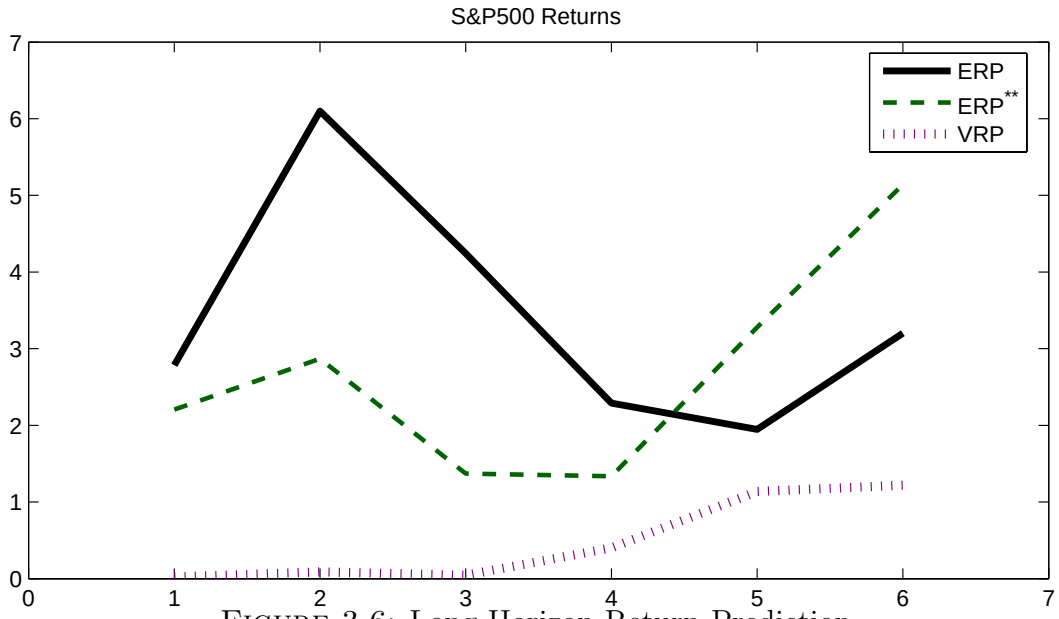


FIGURE 3.6: Long-Horizon Return Prediction

This figure plots the $R^2\%$ in return prediction studies from one to six months. The solid line represents the deeper tail of the jump risk premium $ERPJ(VIX)$ and the diffusive risk premium $ERPV$, the dashed line refers to the benchmark which restricts the shape to be constant with a non self-exciting jump intensity, $X_t = [ERPJ^{**}(VIX), \quad ERPV^{**}]$. The sample ranges from January 1996 to December 2011.

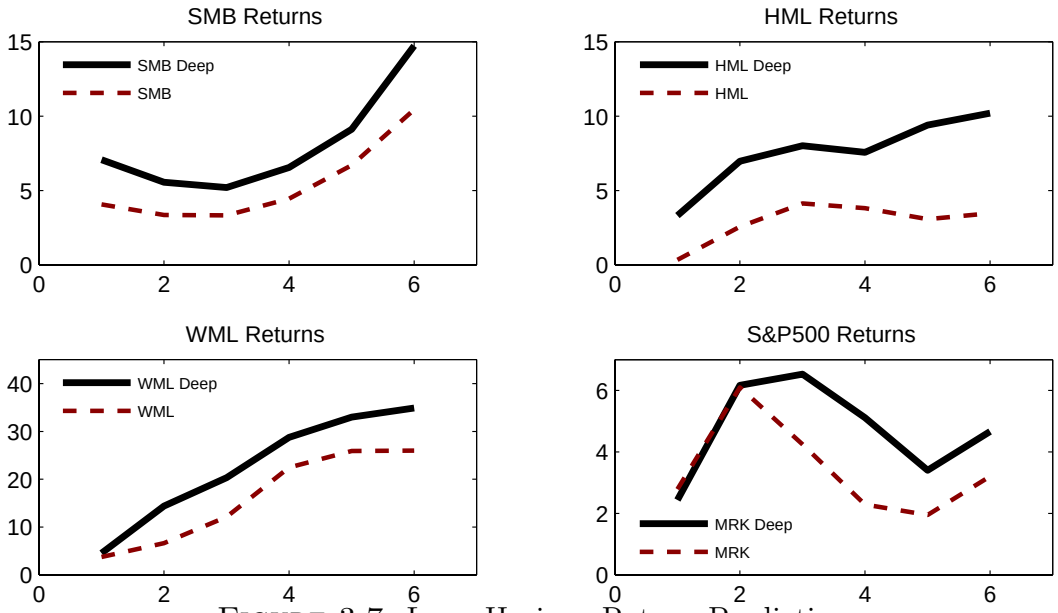


FIGURE 3.7: Long-Horizon Return Prediction

This figure plots the $R^2\%$ for portfolio return prediction studies from one to six months in the top panel. The solid line represents the deeper tail of the portfolio's jump risk premium $ERPJ(VIX)$ and diffusive risk premium $ERPV$, the dashed line refers to the portfolio's total jump risk premium $ERPJ$ and diffusive risk premium $ERPV$. The sample ranges from January 1996 to December 2011.

Appendix A

Supplementary Appendix to "Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence"

This appendix provides additional empirical results pertaining to: (i) summary statistics for the longest available sample for each of the individual countries included in our analysis; (ii) country specific regressions based on the simple VRP^i proxies and the longest available sample for each of the individual countries; (iii) country specific regressions based on the forward $FVRP^i$s discussed in the main text; (iv) country specific regressions based on the $FVRP^{global}$ proxy; (v) panel regressions based on the alternative $FVRP^{global}$ proxy constructed from a standard VAR(1) for the monthly realized and options implied variation measures for all of the six countries; (vi) time series plots of the forward $FVRP^i$s for each of the six individual countries; (vii) plots of the panel regression coefficients and R^2s reported in Table 7, along with the panel regression coefficients and R^2s obtained from regressing the excess returns on the individual country variance risk premia $FVRP^i$; (viii) plots of the R^2s for each of

the individual countries implied by the panel regression results reported in the top panel in Table 7; and (ix) plots of the simple VRP^{global} proxy and the alternative "global" $FVRP$ constructed from a standard VAR(1) for the monthly realized and options implied variation measures for all of the six countries.

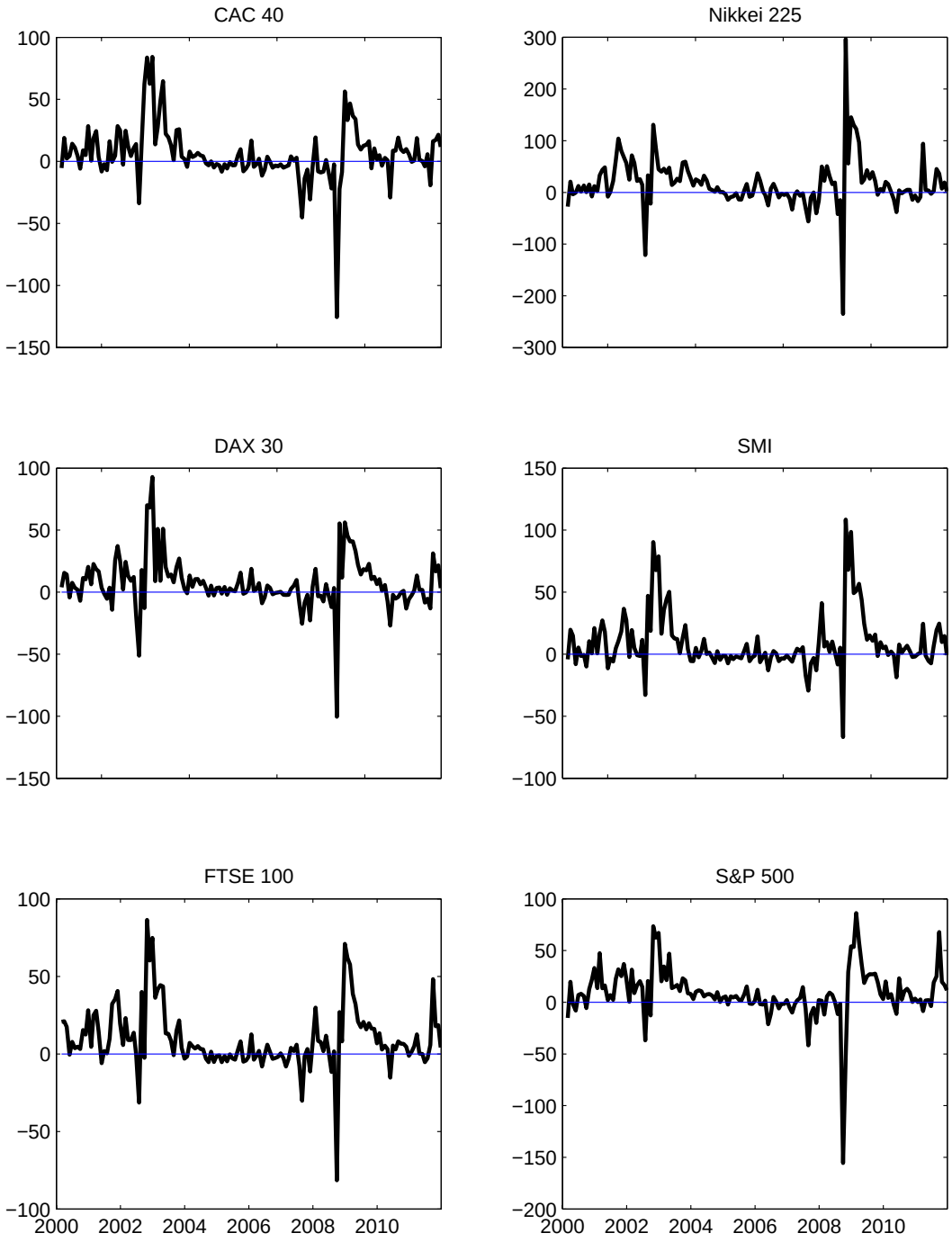


FIGURE A.1: Forward Variance Risk Premia

The figure shows the monthly forward variance risk premia $FVRP_t^i$ for France (CAC 40), Japan (Nikkei 225), Germany (DAX 30), Switzerland (SMI 20), the U.K. (FTSE 100), and the U.S. (S&P 500). The risk premia are constructed by subtracting the HAR-RV-based forecasts from the model-free options implied variation. The sample period spans January 2000 to December 2011

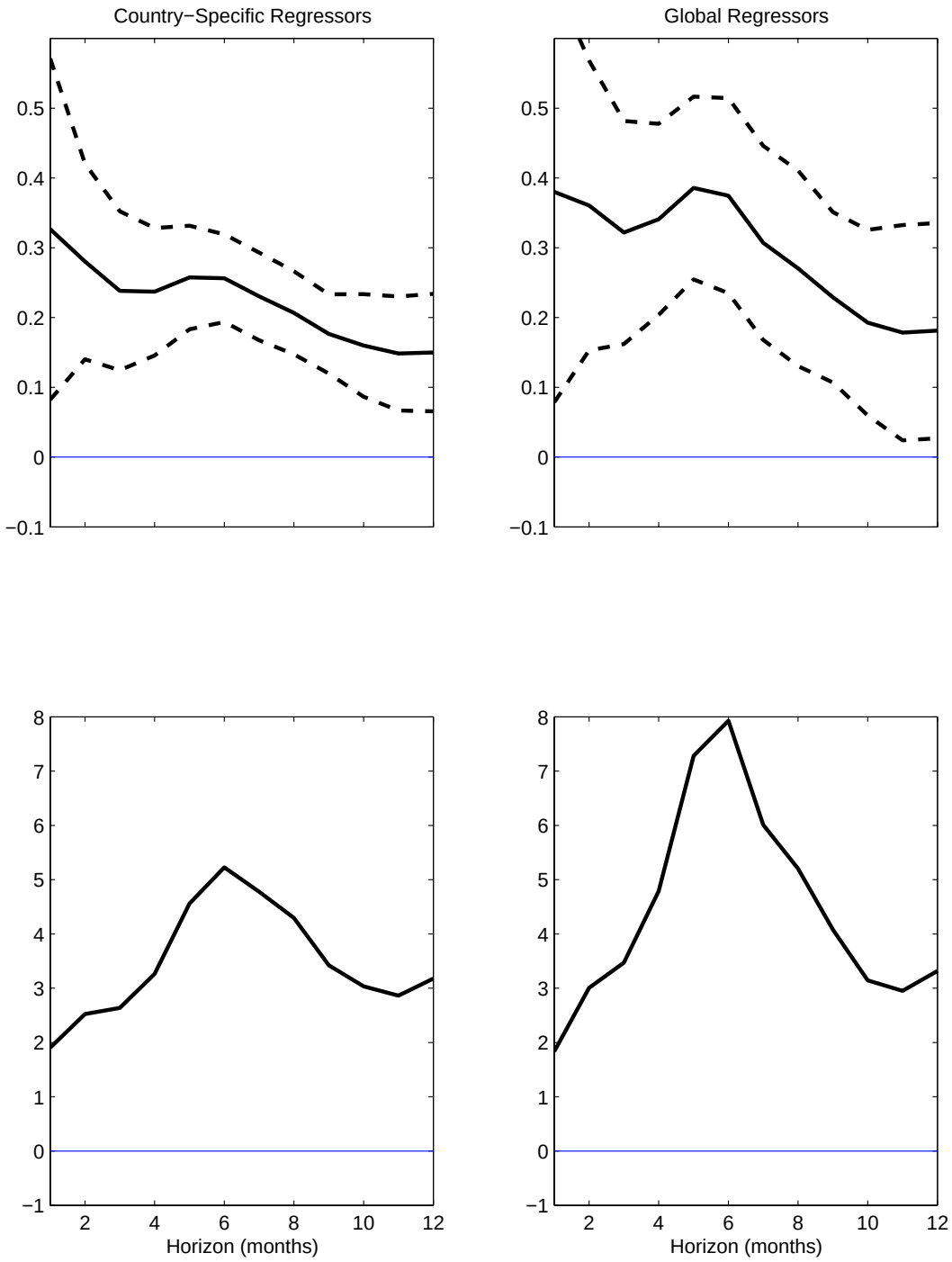


FIGURE A.2: FVRP Panel Regression Coefficients and R^2's

The two top panels show the estimated panel regression coefficients from regressing the excess returns on the individual country variance risk premia $FVRP_t^i$ and the "global' variance risk premium $FVRP_t^{global}$, respectively, together with two NW-based standard error bands. The bottom two panels show the $R^2(h)$'s from the same two panel regressions. The regressions are based on monthly data from January 2000 through December 2011.

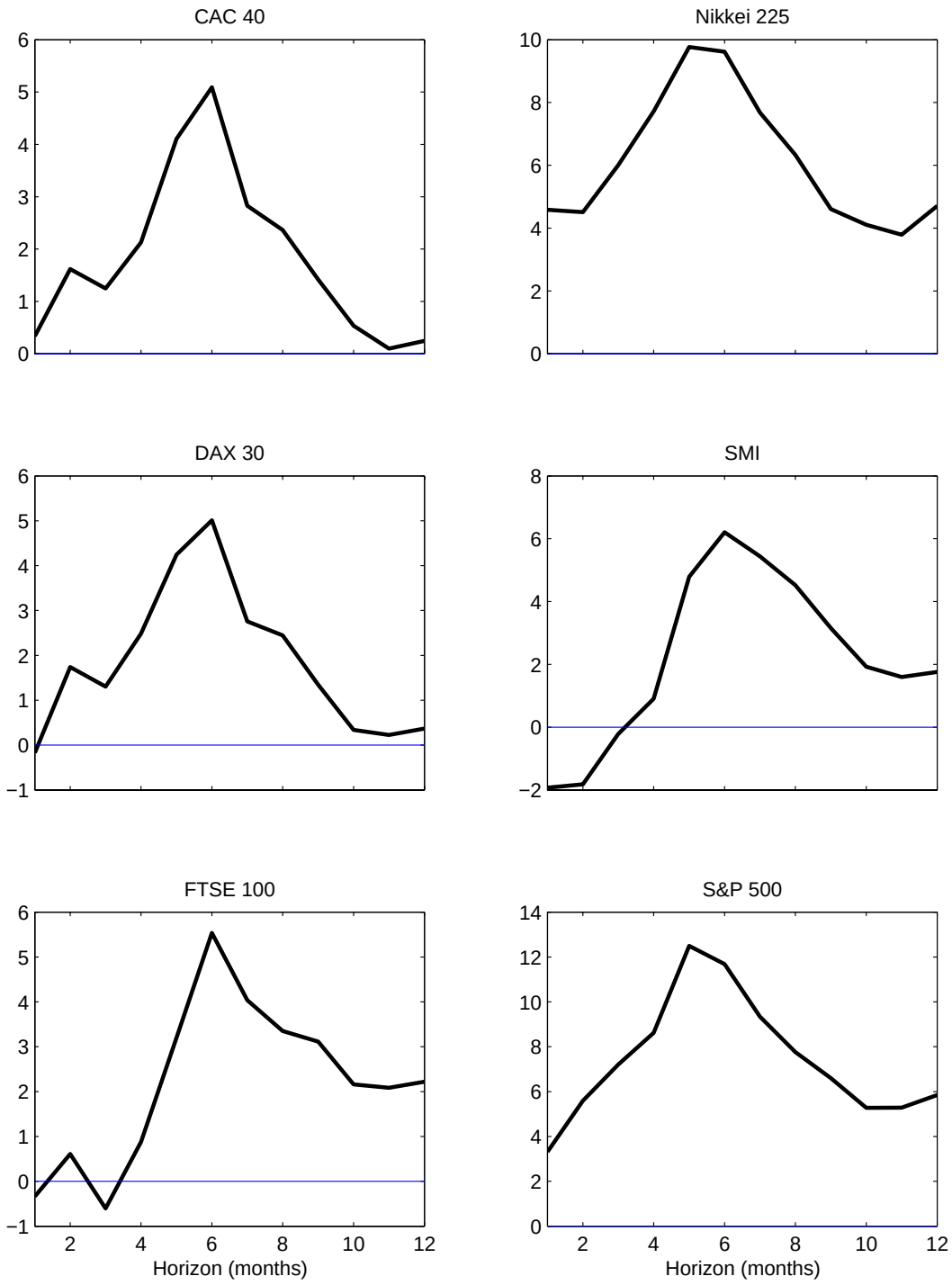


FIGURE A.3: $FVRP^{global}$ Panel Regression Implied R^2's

The figure shows the adjusted $R^2(h)$'s for each of the individual countries implied by the $FVRP_t^{global}$ panel regressions reported in the top panel in Table 7. The regressions are based on monthly data from January 2000 to December 2011.

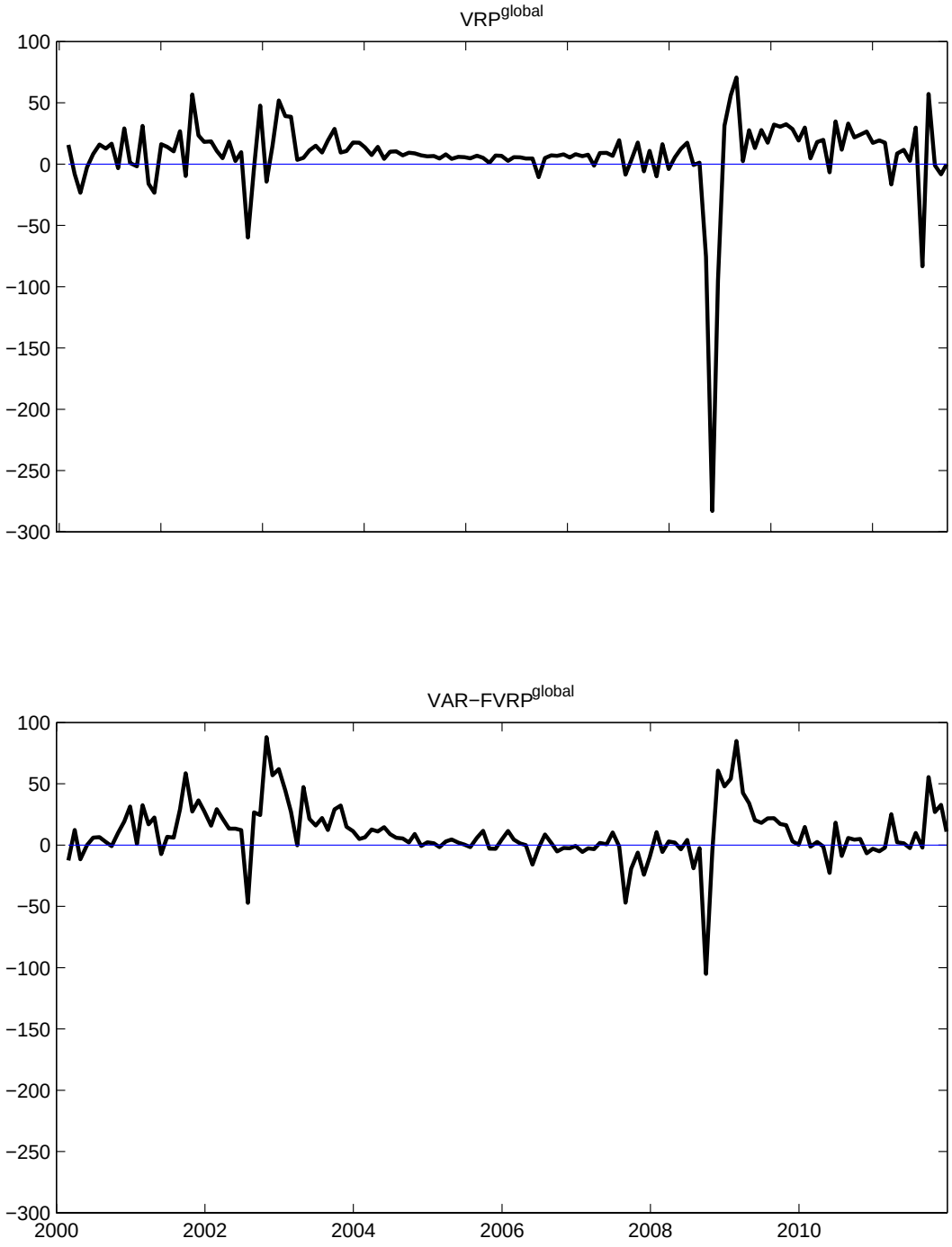


FIGURE A.4: VAR-based $FVRP^{global}$

The figure shows the simple variance risk premia VRP^{global} (top panel) and the forward VAR-based $FVRP^{global}$ (bottom panel). The sample period spans from January 2000 to December 2011.

Table A.1: Summary Statistics for Extended Samples

The monthly excess returns are in annualized percentage form. The variance risk premia are in monthly percentage-squared form. The sample periods for each of the different indexes are given in the second row in the table.

	CAC 40		DAX 30		FTSE 100		Nikkei 225		SMI		S&P 500		Global
Sample	2000.01-2011.12		1994.12-2011.12		2000.01-2011.12		1998.01-2011.12		1999.01-2011.12		1990.01-2011.12		2000.01-2011.12
	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	$r_t - r_{f,t}$	VRP_t	VRP_t
Mean	-8.35	2.75	3.24	6.57	-4.69	7.55	-5.16	12.78	-2.32	7.55	2.73	10.73	7.54
Std. Dev	67.44	41.69	81.37	30.29	52.32	32.16	70.95	37.16	51.16	27.34	52.95	27.42	32.59
Skewness	-0.58	-4.84	-0.87	-2.83	-0.63	-5.56	-0.76	-4.61	-0.73	-3.35	-0.74	-5.74	-5.41
Kurtosis	3.60	41.98	5.33	17.57	3.55	50.72	4.62	42.43	3.53	25.05	4.48	56.45	47.00
AR(1)	0.13	0.30	0.05	0.10	0.06	0.34	0.11	0.14	0.25	0.16	0.07	0.39	0.36

Table A.2: Country Specific Regressions based on Extended Samples

The results are based on the country specific regressions in equation (2) and the longest available sample for each of the individual countries, as indicated in the previous table. t^{NW}-statistics are reported in parentheses.

Index	Horizon	1	2	3	4	5	6	9	12
CAC 40	Constant	-9.00	-9.09	-9.17	-9.49	-9.47	-9.13	-8.33	-7.93
		(-1.50)	(-1.51)	(-1.52)	(-1.54)	(-1.49)	(-1.42)	(-1.24)	(-1.15)
	VRP_t^i	0.24	0.22	0.22	0.22	0.18	0.13	0.05	0.04
		(2.18)	(2.63)	(3.45)	(4.43)	(4.66)	(3.19)	(1.41)	(0.90)
	$Adj.R^2$	1.42	2.66	3.99	5.35	3.98	2.15	-0.15	-0.42
DAX 30	Constant	3.58	2.67	2.58	1.98	1.95	2.76	3.54	3.44
		(0.53)	(0.40)	(0.41)	(0.32)	(0.31)	(0.44)	(0.55)	(0.53)
	VRP_t^i	-0.05	0.12	0.13	0.21	0.21	0.12	0.06	0.09
		(-0.29)	(0.67)	(1.03)	(2.01)	(2.55)	(1.69)	(1.44)	(2.13)
	$Adj.R^2$	-0.46	-0.13	0.15	1.67	1.86	0.49	-0.16	0.51
FTSE 100	Constant	-4.96	-5.44	-6.02	-6.43	-6.42	-6.09	-5.32	-5.05
		(-1.03)	(-1.15)	(-1.33)	(-1.40)	(-1.36)	(-1.27)	(-1.07)	(-0.98)
	VRP_t^i	0.04	0.07	0.14	0.17	0.16	0.13	0.04	0.01
		(0.35)	(0.95)	(3.12)	(4.09)	(3.95)	(2.31)	(1.01)	(0.29)
	$Adj.R^2$	-0.67	-0.40	1.43	3.44	3.18	2.05	-0.39	-0.74
Nikkei 225	Constant	-4.05	-5.14	-5.92	-6.08	-5.62	-5.28	-4.13	-3.92
		(-0.64)	(-0.88)	(-0.99)	(-1.03)	(-0.96)	(-0.90)	(-0.69)	(-0.65)
	VRP_t^i	-0.09	0.00	0.08	0.10	0.07	0.05	0.01	0.01
		(-0.83)	(0.01)	(0.96)	(1.27)	(0.98)	(0.78)	(0.13)	(0.25)
	$Adj.R^2$	-0.40	-0.61	-0.12	0.32	-0.10	-0.28	-0.63	-0.64
SMI	Constant	-2.43	-3.34	-3.67	-4.40	-4.71	-4.36	-3.50	-3.38
		(-0.43)	(-0.60)	(-0.67)	(-0.82)	(-0.88)	(-0.82)	(-0.64)	(-0.60)
	VRP_t^i	0.02	0.11	0.14	0.23	0.25	0.21	0.14	0.14
		(0.11)	(1.01)	(1.40)	(2.24)	(3.15)	(2.80)	(2.57)	(3.25)
	$Adj.R^2$	-0.65	-0.12	0.51	3.18	4.37	3.53	1.88	2.25
S&P 500	Constant	-2.76	-1.71	-1.82	-1.57	-1.30	-0.22	1.54	2.02
		(-0.90)	(-0.54)	(-0.61)	(-0.51)	(-0.41)	(-0.07)	(0.46)	(0.58)
	VRP_t^i	0.51	0.40	0.41	0.38	0.33	0.24	0.11	0.07
		(5.00)	(4.85)	(7.60)	(6.97)	(5.77)	(4.01)	(2.15)	(1.45)
	$Adj.R^2$	6.65	7.71	12.24	13.03	11.06	6.48	1.64	0.70

Table A.3: Country Specific FVRP Regressions

The results are based on the country specific regressions in equation (2) with the forward variance risk premia $FVRP^i$ for each of the countries in place of VRP^i. t^{NW}-statistics are reported in parentheses. The sample period extends from January 2000 to December 2011.

Index	Horizon	1	2	3	4	5	6	9	12
CAC 40	Constant	-11.82	-11.01	-10.54	-10.96	-10.95	-10.66	-9.66	-9.12
		(-1.75)	(-1.62)	(-1.58)	(-1.63)	(-1.61)	(-1.58)	(-1.41)	(-1.29)
	$FVRP_t^i$	0.60	0.46	0.37	0.41	0.39	0.36	0.26	0.23
		(2.55)	(2.08)	(1.49)	(1.59)	(1.63)	(1.91)	(1.91)	(2.10)
	$Adj.R^2$	3.18	3.32	3.22	5.27	5.63	5.52	3.79	3.64
DAX 30	Constant	-5.59	-7.05	-6.21	-6.55	-7.01	-7.20	-5.72	-5.23
		(-0.65)	(-0.84)	(-0.76)	(-0.82)	(-0.86)	(-0.89)	(-0.72)	(-0.65)
	$FVRP_t^i$	0.11	0.33	0.24	0.27	0.34	0.40	0.27	0.26
		(0.27)	(0.94)	(0.82)	(1.09)	(1.49)	(2.15)	(1.67)	(2.10)
	$Adj.R^2$	-0.63	0.54	0.22	0.87	2.19	3.90	2.23	2.63
FTSE 100	Constant	-6.03	-6.50	-6.39	-6.90	-7.21	-7.50	-6.55	-6.21
		(-1.05)	(-1.17)	(-1.19)	(-1.29)	(-1.33)	(-1.39)	(-1.24)	(-1.16)
	$FVRP_t^i$	0.14	0.16	0.15	0.19	0.22	0.26	0.16	0.13
		(0.46)	(0.74)	(0.80)	(1.09)	(1.24)	(1.70)	(1.00)	(0.98)
	$Adj.R^2$	-0.43	0.02	0.25	1.21	2.20	3.72	1.54	1.13
Nikkei 225	Constant	-13.82	-12.39	-11.02	-10.03	-9.87	-9.39	-7.72	-6.72
		(-1.68)	(-1.63)	(-1.53)	(-1.42)	(-1.39)	(-1.34)	(-1.12)	(-0.96)
	$FVRP_t^i$	0.38	0.29	0.23	0.19	0.20	0.19	0.12	0.09
		(2.82)	(3.35)	(3.34)	(2.91)	(2.97)	(3.14)	(2.30)	(1.84)
	$Adj.R^2$	5.76	5.85	5.27	4.52	5.55	6.09	3.34	2.48
SMI	Constant	-1.50	-2.16	-2.78	-3.45	-4.46	-5.05	-5.26	-5.06
		(-0.26)	(-0.36)	(-0.46)	(-0.58)	(-0.72)	(-0.80)	(-0.81)	(-0.76)
	$FVRP_t^i$	-0.11	-0.09	-0.03	0.01	0.11	0.17	0.20	0.18
		(-0.60)	(-0.48)	(-0.20)	(0.10)	(0.85)	(1.45)	(1.77)	(1.79)
	$Adj.R^2$	-0.48	-0.47	-0.68	-0.72	0.00	1.13	2.75	2.94
S&P 500	Constant	-8.62	-8.26	-7.99	-8.08	-8.28	-7.59	-6.07	-5.54
		(-1.33)	(-1.42)	(-1.46)	(-1.47)	(-1.53)	(-1.47)	(-1.17)	(-1.04)
	$FVRP_t^i$	0.61	0.52	0.49	0.50	0.52	0.45	0.28	0.24
		(2.04)	(2.55)	(2.86)	(2.68)	(3.19)	(3.76)	(2.67)	(2.71)
	$Adj.R^2$	5.94	7.89	10.36	13.10	16.79	14.23	7.52	6.82

Table A.4: "Global" FVRP Regressions

The results are based on the country specific regressions in equation (3) with the forward $FVRP^{global}$ in place of VRP^{global}. t^{NW}-statistics are reported in parentheses. The sample period extends from January 2000 to December 2011.

Index	Horizon	1	2	3	4	5	6	9	12
CAC 40	Constant	-11.61	-11.88	-11.24	-11.59	-12.05	-11.98	-11.10	-9.42
		(-1.59)	(-1.64)	(-1.59)	(-1.64)	(-1.70)	(-1.73)	(-1.62)	(-1.33)
	$FVRP_t^{global}$	0.32	0.35	0.28	0.30	0.35	0.36	0.28	0.17
		(1.20)	(1.62)	(1.38)	(1.39)	(1.74)	(2.23)	(1.81)	(1.36)
	$Adj.R^2$	0.61	2.00	1.89	2.98	5.08	6.17	4.01	1.86
DAX 30	Constant	-7.49	-8.63	-7.47	-7.85	-8.20	-7.93	-6.78	-4.70
		(-0.90)	(-1.04)	(-0.93)	(-0.97)	(-1.00)	(-0.98)	(-0.85)	(-0.59)
	$FVRP_t^{global}$	0.27	0.40	0.30	0.34	0.39	0.39	0.29	0.15
		(0.76)	(1.48)	(1.25)	(1.38)	(1.75)	(2.33)	(1.76)	(0.98)
	$Adj.R^2$	-0.08	1.78	1.34	2.51	4.29	5.09	2.87	0.69
FTSE 100	Constant	-6.84	-7.21	-6.57	-7.07	-7.57	-7.71	-7.28	-6.36
		(-1.22)	(-1.39)	(-1.31)	(-1.41)	(-1.49)	(-1.55)	(-1.47)	(-1.24)
	$FVRP_t^{global}$	0.21	0.23	0.16	0.20	0.25	0.27	0.23	0.14
		(0.83)	(1.29)	(1.07)	(1.42)	(1.78)	(2.40)	(1.87)	(1.44)
	$Adj.R^2$	0.25	1.32	0.87	2.33	4.74	6.52	4.75	2.51
Nikkei 225	Constant	-15.76	-13.16	-12.55	-12.02	-11.78	-10.86	-10.00	-7.63
		(-1.78)	(-1.73)	(-1.79)	(-1.72)	(-1.71)	(-1.61)	(-1.48)	(-1.10)
	$FVRP_t^{global}$	0.83	0.56	0.55	0.54	0.54	0.48	0.40	0.25
		(2.20)	(2.39)	(3.27)	(2.86)	(3.08)	(3.51)	(3.47)	(2.38)
	$Adj.R^2$	6.93	5.47	7.57	9.23	10.80	10.23	8.26	5.16
SMI	Constant	-3.80	-4.45	-4.98	-5.50	-6.50	-6.68	-6.38	-4.96
		(-0.61)	(-0.72)	(-0.80)	(-0.89)	(-1.04)	(-1.08)	(-1.03)	(-0.79)
	$FVRP_t^{global}$	0.13	0.15	0.19	0.23	0.32	0.33	0.29	0.15
		(0.60)	(0.75)	(1.02)	(1.22)	(1.84)	(2.40)	(2.20)	(1.38)
	$Adj.R^2$	-0.35	0.02	0.90	1.97	5.47	6.72	5.79	2.16
S&P 500	Constant	-7.99	-8.14	-7.89	-7.79	-8.21	-7.70	-7.14	-5.71
		(-1.19)	(-1.32)	(-1.36)	(-1.34)	(-1.41)	(-1.39)	(-1.31)	(-1.05)
	$FVRP_t^{global}$	0.51	0.48	0.45	0.43	0.47	0.42	0.36	0.24
		(1.75)	(2.33)	(2.63)	(2.38)	(2.71)	(3.13)	(2.79)	(2.53)
	$Adj.R^2$	3.75	6.21	8.16	9.33	13.26	12.17	9.88	6.61

Table A.5: Panel Regressions with VAR-based FVRP

The results are based on the forward "global" variance risk premium proxy constructed from a VAR(1) for the monthly realized and options implied variation measures for all of the six countries. NW-based t-statistics are reported in parentheses. The sample period extends from January 2000 to December 2011.

Horizon	1	2	3	4	5	6	9	12
Constant	-9.38	-9.06	-8.45	-8.48	-8.97	-8.86	-7.46	-6.72
	(-2.71)	(-2.97)	(-2.93)	(-3.06)	(-3.38)	(-3.61)	(-3.66)	(-3.47)
$VAR - FVRP_t^{global}$	0.44	0.39	0.34	0.34	0.39	0.39	0.27	0.22
	(2.65)	(3.37)	(3.59)	(4.31)	(5.09)	(5.07)	(4.13)	(2.71)
$Adj.R^2$	2.22	3.15	3.38	4.22	6.55	7.68	5.01	4.18
Constant	12.20	8.23	8.13	15.68	23.22	26.36	22.69	12.49
	(0.44)	(0.33)	(0.36)	(0.75)	(1.11)	(1.31)	(1.52)	(1.36)
$VAR - FVRP_t^{global}$	0.47	0.42	0.36	0.38	0.45	0.45	0.32	0.25
	(3.17)	(4.32)	(4.73)	(6.30)	(7.80)	(7.88)	(7.39)	(4.17)
$\log(P_t/E_t)^{global}$	-9.98	-7.99	-7.65	-11.15	-14.83	-16.21	-13.83	-8.79
	(-0.77)	(-0.68)	(-0.72)	(-1.17)	(-1.56)	(-1.77)	(-2.08)	(-2.30)
$Adj.R^2$	2.28	3.23	3.53	4.77	7.80	9.43	6.75	5.03

Appendix B

Supplementary Appendix to "Stock Return and Cash Flow Predictability: The Role of Volatility Risk"

B.1 Model Solution

Our basic solution method for the model is adopted from Bansal and Yaron (2004a), Bansal, Kiku, and Yaron (2007a), and Drechsler and Yaron (2011b). To begin, we follow Campbell and Shiller (1988c) and solve for the return on consumption by log-linearizing $r_{c,t+1}$ around the unconditional mean of the wealth-consumption ratio ν_t,

$$r_{c,t+1} \approx \kappa_0 + \kappa_1 \nu_{t+1} - \nu_t + \Delta c_{t+1}, \tag{B.1}$$

where $\kappa_1 = \frac{\exp(E(\nu))}{1+\exp(E(\nu))}$, and $\kappa_0 = \log[1 + \exp(E(\nu))] - \kappa_1 E(\nu)$. We then conjecture a solution for ν_t as a linear function of the state vector Y_t,

$$\nu_t = A_0 + A' Y_t, \tag{B.2}$$

where A_0 is a scalar, and $A = (0, A_x, A_\sigma, A_q, 0)$ refer to the pricing coefficients. Next, by substituting ν_t and ν_{t+1} into equation (B.1), both $r_{c,t+1}$ and the stochastic discount

factor m_{t+1} defined in the main text may be expressed as linear functions of the state vector,

$$m_{t+1} = \mu_m - (\gamma e_1' + (1-\theta)\kappa_1 A')Y_{t+1} - (\theta - 1)A'Y_t, \tag{B.3}$$

$$r_{c,t+1} = \mu_{r_c} + (e_1' + \kappa_1 A')Y_{t+1} - A'Y_t. \tag{B.4}$$

Going one step further, it follows that the innovations to the pricing kernel and the return on the wealth claim may be expressed as,

$$m_{t+1} - E_t(m_{t+1}) = \quad -\Lambda' H G_t z_{t+1}, \tag{B.5}$$

$$r_{c,t+1} - E_t(r_{c,t+1}) = \quad \Lambda_c' H G_t z_{t+1}, \tag{B.6}$$

where Λ denotes the price of risk for the factor shocks,

$$\Lambda = \gamma e_1 + \kappa_1(1-\theta)A,$$

for $e_1 \equiv [1,0,0,0,0]$, and $\Lambda_c = e_1 + \kappa_1 A$. The magnitude and sign of Λ are determined by the preference parameter θ and the pricing coefficient vector A. If investors prefer early resolution of uncertainty, i.e., $\gamma > \phi^{-1}$, A reveals the sensitivity of the market prices for the different shocks to higher order consumption dynamics. When $\gamma = \phi^{-1}$ (CRRA case), Λ collapses to γe_1, and only transient shocks to consumption growth level $z_{g,t+1}$ are priced.

Since the no-arbitrage condition must hold regardless of the realization of the state vector Y_t, it is possible to solve for A by imposing the Euler equation,

$$0 = \mu_m + \mu_{r_c} + [(-\Lambda + \Lambda_c)'F - \theta A']Y_t + \frac{1}{2}(-\Lambda + \Lambda_c)' H G_t G_t' H' (-\Lambda + \Lambda_c). \tag{B.7}$$

This in turn implies that

$$\theta A_{\{i\}} + (\tilde{\Lambda}_c' F)_{\{i\}} = \frac{1}{2}[1_{i=3} \sum_{j=1,5} (\tilde{\Lambda}_c' h_j)^2 + 1_{i=4} \sum_{j=2,3,4} (\tilde{\Lambda}_c' h_j)^2], \tag{B.8}$$

$$0 = \mu_m + \mu_{r_c}, \tag{B.9}$$

where $\tilde{\Lambda}_c = -\Lambda_c + \Lambda = (\gamma - 1)e_1 - \kappa_1 \theta A$, i refers to the i^{th} element of vector, and $1_{i=n}$ is an indicator function. The solutions are,

$$A_x = -\frac{\gamma - 1}{\theta(1 - \kappa_1 \rho_x)}, \tag{B.10}$$

$$A_\sigma = \frac{(\gamma - 1)^2}{2\theta(1 - \kappa_1 \rho_\sigma)}, \tag{B.11}$$

while A_q solves the equation $\frac{1}{2} a_q \theta^2 A_q^2 + (b_q + (1 - \kappa_1 \rho_q))(-\theta A_q) + \frac{1}{2} c_q = 0$, where

$$\begin{aligned} a_q &= \kappa_1^2(\varphi_x^2 s_{q,x}^2 + s_{q,\sigma}^2 + \varphi_q^2) > 0, \\ b_q &= \kappa_1^2 \left(\varphi_x^2(-A_x\theta - A_\sigma \theta s_{\sigma,x}) s_{q,x} - A_\sigma \theta s_{q,\sigma}\right), \\ c_q &= \kappa_1^2 \left(\varphi_x^2(-A_x\theta - A_\sigma \theta s_{\sigma,x})^2 + A_\sigma^2 \theta^2\right) > 0. \end{aligned}$$

Since $a_q > 0$ and $c_q > 0$, the two roots are either negative or positive. We choose the larger root for $-\theta A_q$ if $b_q + (1 - \kappa_1 \rho_q) > 0$, or the smaller root if $b_q + (1 - \kappa_1 \rho_q) < 0$. In both cases A_q reduces to zero when $s_{q,x}$, $s_{q,\sigma}$ and φ_q are zero.

Even though no closed-form expressions for A are available when we consider κ_0 and κ_1 as endogenous, the system of equations are still solvable. As shown in equation (B.8), A depends on κ_1, μ, F, H, as well as the preference parameters. Considering the definitions of κ_1 and κ_0, κ_1 and A are the only unknowns in the constant term in the Euler equation, so that κ_1 may be solved endogenously together with A. Finally, κ_0 and A_0 can be expressed as functions of A and κ_1. For detailed numerical solutions, see Drechsler and Yaron (2011b) Appendix A.1 and A.2.

Applying a similar conjecture-evaluation type method, it is possible to solve for

the aggregate market return $r_{t,t+1}$. Denote the price-dividend ratio by w_t, and consider the conjecture solution w_t=$A_{d,0} + A_d' Y_t$. Log-linearize $r_{t,t+1}$ around the unconditional mean of the price-dividend ratio yields,

$$r_{t,t+1} \approx \kappa_{d,0} + \kappa_{d,1} w_{t+1} - w_t + \Delta d_{t+1}. \tag{B.12}$$

Substituting out w_t and w_{t+1} in the above equation, the return on the market may be rewritten as,

$$r_{t,t+1} = \mu_{r_d} + (e_5' + \kappa_{d,1} A_d') Y_{t+1} - A_d' Y_t, \tag{B.13}$$

where Λ_d=$e_5 + \kappa_{d,1} A_d$ and A_d=$[0, A_{d,x}, A_{d,\sigma}, A_{d,q}, A_{d,d}]'$ is a vector of pricing coefficients.

Using the same solution method as the one previously used for A, it follows by the no-arbitrage condition,

$$0 = \mu_m + \mu_{r_d} + [(-\Lambda + \Lambda_d)' F - (\theta - 1) A' - A_d']' Y_t + 0.5(-\Lambda + \Lambda_d)' H G_t G_t' H' (-\Lambda + \Lambda_d), \tag{B.14}$$

which implies that

$$(\theta - 1) A_{\{i\}} + A_{d,\{i\}} + (\tilde{\Lambda}_d' F)_{\{i\}} = 0.5[1_{i=3} \sum_{j=1,5} (\tilde{\Lambda}_d' h_j)^2 + 1_{i=4} \sum_{j=2,3,4} (\tilde{\Lambda}_d' h_j)^2], \tag{B.15}$$

$$0 = \mu_m + \mu_{r_d}, \tag{B.16}$$

where $\tilde{\Lambda}_d$=$-\Lambda_d + \Lambda$=$\gamma e_1 - e_5 + \kappa_1 (1 - \theta) A - \kappa_{d,1} A_d$.

The solution for A_d may therefore be expressed as,

$$A_{d,d} = \frac{\rho_d}{1-\kappa_{d,1}\rho_d} \tag{B.17}$$

$$\frac{1-\kappa_1\rho_x}{1-\kappa_{d,1}\rho_x}(1-\theta)A_x - A_{d,x} = -\frac{-\gamma+\phi_{dx}(1+\kappa_{d,1}A_{d,d})}{1-\kappa_{d,1}\rho_x} \tag{B.18}$$

$$\frac{1-\kappa_1\rho_\sigma}{1-\kappa_{d,1}\rho_\sigma}(1-\theta)A_\sigma - A_{d,\sigma} = -\frac{1}{2}\frac{\gamma^2+\varphi_d^2(1+\kappa_{d,1}A_{d,d})^2}{1-\kappa_{d,1}\rho_\sigma} < 0 \tag{B.19}$$

$$\frac{1-\kappa_1\rho_q}{1-\kappa_{d,1}\rho_q}(1-\theta)A_q - A_{d,q} = -\frac{1}{2}\frac{d_{d,q}}{1-\kappa_{d,1}\rho_q} \tag{B.20}$$

where

$$a_{d,q} = a_q = (\varphi_x^2 s_{q,x}^2 + s_{q,\sigma}^2 + \varphi_q^2) > 0$$

$$b_{d,q} = \varphi_x^2 s_{q,x}[\kappa_1(1-\theta)A_x - \kappa_{d,1}A_{d,x} + \kappa_1(1-\theta)A_\sigma - \kappa_{d,1}A_{d,\sigma}s_{\sigma,x} + \frac{-1}{1-\kappa_{d,1}\rho_d}s_{d,x}]$$
$$+\left(\kappa_1(1-\theta)A_\sigma - \kappa_{d,1}A_{d,\sigma} + \frac{-1}{1-\kappa_{d,1}\rho_d}s_{d,\sigma}\right)s_{q,\sigma} + \frac{-1}{1-\kappa_{d,1}\rho_d}s_{d,q}\varphi_q^2$$

$$c_{d,q} = \varphi_x^2[\kappa_1(1-\theta)A_x - \kappa_{d,1}A_{d,x} + (\kappa_1(1-\theta)A_\sigma - \kappa_{d,1}A_{d,\sigma})\,s_{\sigma,x} + \frac{-1}{1-\kappa_{d,1}\rho_d}s_{d,x}]^2$$
$$+[\left(\kappa_1(1-\theta)A_\sigma - \kappa_{d,1}A_{d,\sigma} + \frac{-1}{1-\kappa_{d,1}\rho_d}s_{d,\sigma}\right)^2 + \frac{-1}{1-\kappa_{d,1}\rho_d}s_{d,q}]^2\varphi_q^2 > 0$$

$$d_{d,q} = a_{d,q}\left(\kappa_1(1-\theta)A_q - \kappa_{d,1}A_{d,q}\right)^2 + 2b_{d,q}\left(\kappa_1(1-\theta)A_q - \kappa_{d,1}A_{d,q}\right) + c_{d,q}$$

In other words, $A_{d,q}$ solves the equation (A.20) and we choose the root with smaller absolute value. Both A_q and $A_{d,q}$ reduce to zero when $s_{q,x}$, $s_{q,\sigma}$ and φ_q are all zero. We will discuss the sign of $A_{d,q}$ later on in the parameter implication section. We

can explicitly express $A_{d,x}$ and $A_{d,\sigma}$ as

$$A_{d,x} = \frac{(1-\gamma)/\theta - 1 + \phi_{dx}/(1-\kappa_{d,1}\rho_d)}{1-\kappa_{d,1}\rho_x} = \frac{-\psi^{-1} + \phi_{dx}/(1-\kappa_{d,1}\rho_d)}{1-\kappa_{d,1}\rho_x} \tag{B.21}$$

$$A_{d,\sigma} = \frac{(\gamma-1)^2 + 2\theta\gamma + \theta(\varphi_d^2/(1-\kappa_{d,1}\rho_d)^2 - 1)}{2\theta(1-\kappa_{d,1}\rho_\sigma)}. \tag{B.22}$$

B.2 Variance Risk Premium

In order to determine the factor structure for the variance risk premium, we first need to solve for the second order moment of the return $r_{t,t+1}$. It follows from above that $r_{t,t+1} - E_t(r_{t,t+1}) = \Lambda_d' H G_t z_{t+1}$, so that the conditional variance of the return is affine in σ_t^2 and q_t,

$$\begin{aligned} Var_t(r_{t,t+1}) = & \quad \textstyle\sum_{j=1,5} \Lambda_d' h_j h_j' \Lambda_d \sigma_t^2 + \sum_{j=2,3,4} \Lambda_d' h_j h_j' \Lambda_d q_t \\ = & \quad (1+\kappa_{d,1}A_{d,d})^2 \varphi_d^2 \sigma_t^2 + \textstyle\sum_{j=2,3,4} \Lambda_d' h_j h_j' \Lambda_d q_t. \end{aligned} \tag{B.23}$$

The first term is associated with the volatility of cash flow shocks, and the second term represents the consumption uncertainty. Accordingly, the equity risk premium may be expressed as,

$$\begin{aligned} \log(E_t R_{t,t+1}) - r_{f,t} = & E_t(r_{t,t+1}) + \frac{1}{2} Var_t(r_{t,t+1}) - r_{f,t} \\ = & - Cov_t(m_{t+1}, r_{t,t+1}) = \sum_{j=2,3,4} \Lambda_d' h_j h_j' \Lambda q_t \end{aligned} \tag{B.24}$$

The first equality comes from the normality distribution of $r_{t,t+1} \equiv \log(R_{t,t+1})$, the second equality comes from the no arbitrage condition and $r_{f,t} = \log(E_t m_{t+1} + \frac{1}{2} Var_t(m_{t+1}))$. The expectations of $Var_t(r_{t,t+1})$ under the physical and risk-neutral

probability measures are,

$$\begin{aligned} E_t(Var_{t+1}(r_{d,t+2})) &= \sum_{j=1,5} \Lambda_d' h_j h_j' \Lambda_d (\mu_\sigma + \rho_\sigma \sigma_t^2) \\ &+ \sum_{j=2,3,4} \Lambda_d' h_j h_j' \Lambda_d (\mu_q + \rho_q q_t), \end{aligned} \tag{B.25}$$

$$\begin{aligned} E_t^Q(Var_{t+1}(r_{d,t+2})) &= \sum_{j=1,5} \Lambda_d' h_j h_j' \Lambda_d (\mu_\sigma + \rho_\sigma \sigma_t^2 + s_{q,1} q_t) \\ &+ \sum_{j=2,3,4} \Lambda_d' h_j h_j' \Lambda_d (\mu_q + \rho_q q_t + s_{q,2} q_t). \end{aligned} \tag{B.26}$$

Under the risk neutral measure, we reweight probabilities according to the pricing kernel $\frac{e^{m_{t,t+1}}}{E_t e^{m_{t,t+1}}}$. If investor prefers early resolution of uncertainty, the shocks z_{t+1}'s conditional mean shifts away from zero. And this shift can be expressed as the conditional covariance between the state vector and SDF $m_{t,t+1}$,

$$\begin{aligned} s_{q,1} q_t &= Cov_t(e_3 H G_t z_{t+1}, -\Lambda' H G_t z_{t+1}) \\ &= -(\varphi_x s_{\sigma,x} h_2' + h_3') \Lambda q_t, \end{aligned} \tag{B.27}$$

$$\begin{aligned} s_{q,2} q_t &= Cov_t(e_4 H G_t z_{t+1}, -\Lambda' H G_t z_{t+1}) \\ &= -(\varphi_x s_{q,x} h_2' + s_{q,\sigma} h_3' + \varphi_q h_4') \Lambda q_t, \end{aligned} \tag{B.28}$$

where,

$$\begin{aligned} s_{q,1} =& -\kappa_1(1-\theta)\left(A_x \varphi_x^2 s_{\sigma,x} + A_\sigma(\varphi_x^2 s_{\sigma,x}^2 + 1) + A_q(\varphi_x^2 s_{\sigma,x} s_{q,x} + s_{q,\sigma})\right) \\ =& -\kappa_1(1-\theta)(\varphi_x^2 s_{\sigma,x} s_{q,x} + s_{q,\sigma}) \left(\frac{A_x \varphi_x^2 s_{\sigma,x} + A_\sigma(\varphi_x^2 s_{\sigma,x}^2 + 1)}{\varphi_x^2 s_{\sigma,x} s_{q,x} + s_{q,\sigma}} + A_q\right) \\ s_{q,2} =& -\kappa_1(1-\theta)\left(A_x \varphi_x^2 s_{q,x} + A_\sigma(\varphi_x^2 s_{\sigma,x} s_{q,x} + s_{q,\sigma}) + A_q(\varphi_x^2 s_{q,x}^2 + s_{q,\sigma}^2 + \varphi_q^2)\right) \\ =& -1/\kappa_1(1-\theta) a_q \left(\frac{b_q}{-\theta a_q} + A_q\right) \end{aligned}$$

By definition, $s_{q,1}$ and $s_{q,2}$ represent the market prices of shocks to σ_t^2 and q_t, respectively. Thus, the variance risk premium is naturally defined by,

$$\begin{aligned} VRP_t \equiv & \quad E_t^Q(Var_{t+1}(r_{d,t+2})) - E_t(Var_{t+1}(r_{d,t+2})) \\ = & \quad (\textstyle\sum_{j=1,5} \Lambda_d' h_j h_j' \Lambda_d s_{q,1} + \sum_{j=2,3,4} \Lambda_d' h_j h_j' \Lambda_d s_{q,2}) q_t. \end{aligned} \tag{B.29}$$

In the main text, we will refer to the expected return variation and the variance risk premium as,

$$\begin{aligned} ERV_t =& \frac{Q_{1,1}}{\rho_\sigma}(\mu_\sigma + \rho_\sigma \sigma_t^2) + \frac{Q_{1,2}}{\rho_q}(\mu_q + \rho_q q_t), \\ VRP_t =& Q_{2,2} q_t \end{aligned}$$

for short, where

$$Q_{1,1} = \quad \textstyle\sum_{j=1,5} \Lambda_d' h_j h_j' \Lambda_d \rho_\sigma > 0, \tag{B.30}$$

$$Q_{1,2} = \quad \textstyle\sum_{j=2,3,4} \Lambda_d' h_j h_j' \Lambda_d \rho_q > 0, \tag{B.31}$$

$$Q_{2,2} = \quad \tfrac{Q_{1,1}}{\rho_\sigma} s_{q,1} + \tfrac{Q_{1,2}}{\rho_q} s_{q,2}. \tag{B.32}$$

In order to determine the signs of $A_{d,x}$, $A_{d,\sigma}$ and $A_{d,q}$, it is informative to write out the formula in terms of the estimated B and $\tilde{\rho}$ matrices,

$$\frac{\phi_{d,x}}{-A_{d,x}} = \tilde{\rho}_{3,4}, \qquad \frac{A_{d,\sigma}}{Q_{1,1}} = B_{4,1}, \qquad \frac{Q_{1,2}}{Q_{2,2}} = -B_{1,2}, \qquad \frac{A_{d,q}}{Q_{2,2}} = B_{4,2} - B_{1,2}B_{4,1}. \tag{B.33}$$

Since $\tilde{\rho}_{3,4} < 0$, $\phi_{d,x}$ and $A_{d,x}$ must have the same signs. Thus, by definition $Q_{1,1} > 0$ and $Q_{1,2} > 0$, which together with the estimates for $B_{4,1} = -0.60 < 0$ and $B_{1,2} = -0.02 < 0$, imply that $A_{d,\sigma} < 0$ and $Q_{2,2} > 0$. Consequently $A_{d,q} = Q_{2,2}(B_{4,2} - B_{1,2}B_{4,1}) = -1.45 Q_{2,2} < 0$.

B.3 Alternative Setups

B.3.1 Separate Volatility Processes

We will consider the following alternative setup for G_t and H, with F unchanged,

$$G_t = \begin{pmatrix} \sigma_t & 0 & 0 & 0 & 0 \\ 0 & \sqrt{q_t} & 0 & 0 & 0 \\ 0 & 0 & \sigma_t & 0 & 0 \\ 0 & 0 & 0 & \sqrt{q_t} & 0 \\ 0 & 0 & 0 & 0 & \sigma_t \end{pmatrix} \qquad H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & s_{\sigma,x} & \varphi_\sigma & 0 & 0 \\ 0 & s_{q,x} & 0 & \varphi_q & 0 \\ 0 & s_{d,x} & \varphi_\sigma s_{d,\sigma} & \varphi_q s_{d,q} & \varphi_d \end{pmatrix} \quad \text{(B.34)}$$

This setup is related to Bansal and Shaliastovich (2013), where the volatilities of x_t and σ_t^2 are modeled as two separate processes.

For simplicity, we use the same general notation as in the main setup for A, Λ, Λ_c and Λ_d. However, the solutions for the pricing coefficients are obviously different from the main setup, except for $A_{d,d}{=}\frac{\rho_d}{1-\kappa_{d,1}\rho_d}$,

$$\theta A_{\{i\}} + (\tilde{\Lambda}_c' F)_{\{i\}} = 0.5[1_{i=3} \sum_{j=1,3,5} (\tilde{\Lambda}_c' h_j)^2 + 1_{i=4} \sum_{j=2,4} (\tilde{\Lambda}_c' h_j)^2], \quad \text{(B.35)}$$

$$(\theta - 1) A_{\{i\}} + A_{d,\{i\}} + (\tilde{\Lambda}_d' F)_{\{i\}} = 0.5[1_{i=3} \sum_{j=1,3,5} (\tilde{\Lambda}_d' h_j)^2 + 1_{i=4} \sum_{j=2,4} (\tilde{\Lambda}_d' h_j)^2]. \quad \text{(B.36)}$$

Since $r_{t,t+1} - E_t(r_{t,t+1}){=}\Lambda_d' H G_t z_{t+1}$, the conditional variance of the return is again affine,

$$Var_t(r_{t,t+1}) = \sum_{j=1,3,5} \Lambda_d' h_j h_j' \Lambda_d \sigma_t^2 + \sum_{j=2,4} \Lambda_d' h_j h_j' \Lambda_d q_t \quad \text{(B.37)}$$

The expectations of $Var_t(r_{t,t+1})$ under the physical and risk-neutral probability mea-

sures may further be expressed as,

$$\begin{aligned} E_t(Var_{t+1}(r_{d,t+2})) &= \sum_{j=1,3,5} \Lambda_d' h_j h_j' \Lambda_d (\mu_\sigma + \rho_\sigma \sigma_t^2) \\ &+ \sum_{j=2,4} \Lambda_d' h_j h_j' \Lambda_d (\mu_q + \rho_q q_t) \end{aligned} \tag{B.38}$$

$$\begin{aligned} E_t^Q(Var_{t+1}(r_{d,t+2})) &= \sum_{j=1,3,5} \Lambda_d' h_j h_j' \Lambda_d (\mu_\sigma + \rho_\sigma \sigma_t^2 + s_{q,1} q_t + s_{\sigma,1} \sigma_t^2) \\ &+ \sum_{j=2,4} \Lambda_d' h_j h_j' \Lambda_d (\mu_q + \rho_q q_t + s_{q,2} q_t) \end{aligned} \tag{B.39}$$

If investors prefer early resolution of uncertainty, the conditional means of the z_{t+1} shocks shift away from zero under the risk-neutral measure,

$$\begin{aligned} s_{\sigma,1}\sigma_t^2 + s_{q,1} q_t &= Cov_t(e_3' H G_t z_{t+1}, -\Lambda' H G_t z_{t+1}) \\ &= -\varphi_\sigma h_3' \Lambda \sigma_t^2 - (\varphi_x s_{\sigma,x} h_2') \Lambda q_t, \end{aligned} \tag{B.40}$$

$$\begin{aligned} s_{q,2} q_t &= Cov_t(e_4 H G_t z_{t+1}, -\Lambda' H G_t z_{t+1}) \\ &= -(\varphi_x s_{q,x} h_2' + \varphi_q h_4') \Lambda q_t. \end{aligned} \tag{B.41}$$

Defining the variance risk premium as before,

$$\begin{aligned} VRP_t \equiv & E_t^Q(Var_{t+1}(r_{d,t+2})) - E_t(Var_{t+1}(r_{d,t+2})) \\ = & \sum_{j=1,3,5} \Lambda_d' h_j h_j' \Lambda_d (s_{\sigma,1}\sigma_t^2 + s_{q,1} q_t) + \sum_{j=2,4} \Lambda_d' h_j h_j' \Lambda_d s_{q,2} q_t, \end{aligned} \tag{B.42}$$

we may express the expected return variation and premium in short-hand form as,

$$\begin{aligned} ERV_t = & \frac{Q_{1,1}}{\rho_\sigma}(\mu_\sigma + \rho_\sigma \sigma_t^2) + \frac{Q_{1,2}}{\rho_q}(\mu_q + \rho_q q_t), \\ VRP_t = & Q_{2,1}\sigma_t^2 + Q_{2,2} q_t, \end{aligned}$$

where

$$Q_{1,1} = \sum_{j=1,3,5} \Lambda_d' h_j h_j' \Lambda_d \rho_\sigma > 0 \qquad Q_{1,2} = \sum_{j=2,4} \Lambda_d' h_j h_j' \Lambda_d \rho_q > 0$$

$$Q_{2,1} = \frac{Q_{1,1}}{\rho_q} s_{\sigma,1} \qquad Q_{2,2} = \frac{Q_{1,1}}{\rho_\sigma} s_{q,1} + \frac{Q_{1,2}}{\rho_q} s_{q,2}.$$

B.3.2 Long-Run Stochastic Volatility

We will consider the following alternative setup for G_t, H, and F,

$$F = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & \rho_x & 0 & 0 & 0 \\ 0 & 0 & \rho_\sigma & 1 & 0 \\ 0 & 0 & 0 & \rho_q & 0 \\ 0 & \phi_{dx} & 0 & 0 & \rho_d \end{pmatrix} \quad G_t = \begin{pmatrix} \sigma_t & 0 & 0 & 0 & 0 \\ 0 & \sigma_t & 0 & 0 & 0 \\ 0 & 0 & \sigma_t & 0 & 0 \\ 0 & 0 & 0 & \sqrt{q_t} & 0 \\ 0 & 0 & 0 & 0 & \sigma_t \end{pmatrix} \tag{B.43}$$

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \varphi_x & 0 & 0 & 0 \\ 0 & \varphi_x s_{\sigma,x} & \varphi_\sigma & 0 & 0 \\ 0 & \varphi_x s_{q,x} & 0 & \varphi_q & 0 \\ 0 & \varphi_x s_{d,x} & \varphi_\sigma s_{d,\sigma} & \varphi_q s_{d,q} & \varphi_d \end{pmatrix} \tag{B.44}$$

This setup is motivated by the model analyzed by Branger and Vòlkert (2012), among others, allowing for a time-varying mean of the consumption variance σ_t^2.

Again, for simplicity we will use the same general notation as in the main setup for A, Λ, Λ_c and Λ_d. The solution for $A_{d,d}$=$\frac{\rho_d}{1-\kappa_{d,1}\rho_d}$ remains the same, but the other the pricing coefficients now take the form,

$$\theta A_{\{i\}} + (\tilde{\Lambda}_c' F)_{\{i\}} = \frac{1}{2}[1_{i=3} \sum_{j=1,2,3,5} (\tilde{\Lambda}_c' h_j)^2 + 1_{i=4} \sum_{j=4} (\tilde{\Lambda}_c' h_j)^2], \tag{B.45}$$

$$(\theta - 1) A_{\{i\}} + A_{d,\{i\}} + (\tilde{\Lambda}_d' F)_{\{i\}} = \frac{1}{2}[1_{i=3} \sum_{j=1,2,3,5} (\tilde{\Lambda}_d' h_j)^2 + 1_{i=4} \sum_{j=4} (\tilde{\Lambda}_d' h_j)^2]. \tag{B.46}$$

As before, $r_{t,t+1} - E_t(r_{t,t+1})$=$\Lambda_d' H G_t z_{t+1}$, so that the conditional variance of the

return may be expressed as,

$$Var_t(r_{t,t+1}) = \sum_{j=1,2,3,5} \Lambda_d' h_j h_j' \Lambda_d \sigma_t^2 + \sum_{j=4} \Lambda_d' h_j h_j' \Lambda_d q_t. \tag{B.47}$$

The expectation of $Var_t(r_{t,t+1})$ under the physical and risk-neutral probability measures are,

$$E_t(Var_{t+1}(r_{d,t+2})) = \sum_{j=1,2,3,5} \Lambda_d' h_j h_j' \Lambda_d(\mu_\sigma + \rho_\sigma \sigma_t^2 + q_t) + \sum_{j=4} \Lambda_d' h_j h_j' \Lambda_d(\mu_q + \rho_q q_t),$$

$$\begin{aligned} E_t^Q(Var_{t+1}(r_{d,t+2})) = & \sum_{j=1,2,3,5} \Lambda_d' h_j h_j' \Lambda_d(\mu_\sigma + \rho_\sigma \sigma_t^2 + q_t + s_{\sigma,1}\sigma_t^2) \\ & + \sum_{j=4} \Lambda_d' h_j h_j' \Lambda_d(\mu_q + \rho_q q_t + s_{\sigma,2}\sigma_t^2 + s_{q,2} q_t). \end{aligned}$$

The shifts in the conditional means of the z_{t+1} shocks under the risk-neutral measure become,

$$\begin{aligned} s_{\sigma,1}\sigma_t^2 = & \; Cov_t(e_3' H G_t z_{t+1}, -\Lambda' H G_t z_{t+1}) \\ = & \; -\varphi_\sigma h_3' \Lambda \sigma_t^2 - (\varphi_x s_{\sigma,x} h_2') \Lambda \sigma_t^2, \end{aligned} \tag{B.48}$$

$$\begin{aligned} s_{\sigma,2}\sigma_t^2 + s_{q,2} q_t = & \; Cov_t(e_4 H G_t z_{t+1}, -\Lambda' H G_t z_{t+1}) \\ = & \; -(\varphi_x s_{q,x} h_2') \Lambda \sigma_t^2 - (\varphi_q h_4') \Lambda q_t. \end{aligned} \tag{B.49}$$

As before, the expected return variation and variance risk premium, may be conveniently expressed as,

$$\begin{aligned} ERV_t = & \frac{Q_{1,1}}{\rho_\sigma}(\mu_\sigma + \rho_\sigma \sigma_t^2) + \frac{Q_{1,2}}{\rho_q}(\mu_q + \rho_q q_t), \\ VRP_t = & Q_{2,1}\sigma_t^2 + Q_{2,2} q_t, \end{aligned}$$

where

$$Q_{1,1} = \sum_{j=1,2,3,5} \Lambda_d' h_j h_j' \Lambda_d \rho_\sigma > 0 \qquad Q_{1,2} = \sum_{j=4} \Lambda_d h_j h_j' \Lambda_d \rho_q > 0$$

$$Q_{2,1} = \frac{Q_{1,1}}{\rho_\sigma} s_{\sigma,1} + \frac{Q_{1,2}}{\rho_q} s_{\sigma,2} \qquad Q_{2,2} = \frac{Q_{1,2}}{\rho_q} s_{q,2}.$$

B.4 Detailed Derivations for Section 2.2

Substituting f_t by $Q^{-1}(X_t - \mu_X)$ in the basic relation $f_{t+1} = \mu + \rho f_t + S\epsilon_{t+1}$, it follows that

$$Q^{-1}X_{t+1} = \mu + Q^{-1}\mu_X - \rho Q^{-1}\mu_X + \rho Q^{-1}X_t + S\epsilon_{t+1}. \tag{B.50}$$

Normalizing each element of $Q^{-1}X_{t+1}$ by the corresponding diagonal element of Q^{-1}, the model may be rewritten as,

$$BX_{t+1} = \tilde{\mu} + \tilde{\rho}BX_t + \tilde{S}\tilde{\epsilon}_{t+1},$$

where

$$B \equiv \left(\frac{1}{\text{diag}(Q^{-1})}l_{1\times 4}\right) \odot Q^{-1}.$$

To match with equation (B.50),

$$\tilde{\mu} = \left(\frac{1}{\text{diag}(Q^{-1})}l_{1\times 4}\right) \odot (\mu - \rho Q^{-1}\mu_X),$$

and

$$\begin{aligned}
\tilde{\rho} &= \left[\left(\frac{1}{\text{diag}(Q^{-1})}l_{1\times 4}\right) \odot (\rho Q^{-1})\right] B^{-1} \\
&= \left[\left(\frac{1}{\text{diag}(Q^{-1})}l_{1\times 4}\right) \odot (Q^{-1}\text{diag}(\rho) + (\rho - \text{diag}(\rho))Q^{-1})\right] B^{-1} \\
&= \left[\left(\frac{1}{\text{diag}(Q^{-1})}l_{1\times 4}\right) \odot (Q^{-1} \odot (vec(\text{diag}(\rho))l_{1\times 4}) + (\rho - \text{diag}(\rho))Q^{-1})\right] B^{-1} \\
&= \left[B \odot (vec(\text{diag}(\rho))l_{1\times 4} + \frac{\rho - \text{diag}(\rho)}{-A_{d,x}}B\right] B^{-1} \\
&= \rho + \frac{\rho - \text{diag}(\rho)}{-A_{d,x}}
\end{aligned} \tag{B.51}$$

or

Defining $\tilde{\epsilon}_{t+1}$ as

$$\tilde{\epsilon}_{t+1} \equiv \frac{1}{\text{diag}(Q^{-1})} \odot \epsilon_{t+1},$$

it follows again from equation (B.50) that

$$\tilde{S} = \left(\frac{1}{\text{diag}(Q^{-1})} l_{1\times 4}\right) \odot S \odot \frac{1}{\frac{1}{\text{diag}_{(Q^{-1})'}} l_{1\times 4}}.$$

Based on the formula for Q in the main text, the inverse Q^{-1} and $\frac{1}{\text{diag}(Q^{-1})}$ may be expressed as,

$$Q^{-1} = \begin{pmatrix} \frac{1}{Q_{1,1}} & \frac{-Q_{1,2}}{Q_{1,1}Q_{2,2}} & 0 & 0 \\ 0 & \frac{1}{Q_{2,2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{-A_{d,\sigma}}{Q_{1,1}A_{d,x}} & \frac{-Q_{1,1}A_{d,q}+Q_{1,2}A_{d,\sigma}}{Q_{1,1}Q_{2,2}A_{d,x}} & \frac{-A_{d,d}}{A_{d,x}} & -\frac{1}{A_{d,x}} \end{pmatrix} \qquad \frac{1}{\text{diag}(Q^{-1})} = \begin{pmatrix} Q_{1,1} \\ Q_{2,2} \\ 1 \\ -A_{d,x} \end{pmatrix}$$

Combining the expressions for ρ and S, it therefore follows that

$$B = \begin{pmatrix} 1 & -\frac{Q_{1,2}}{Q_{2,2}} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{A_{d,\sigma}}{Q_{1,1}} & \frac{Q_{1,1}A_{d,q}-A_{d,\sigma}Q_{1,2}}{Q_{1,1}Q_{2,2}} & A_{d,d} & 1 \end{pmatrix} \qquad \tilde{\rho} = \begin{pmatrix} \rho_\sigma & 0 & 0 & 0 \\ 0 & \rho_q & 0 & 0 \\ 0 & 0 & \rho_d & \frac{\phi_{dx}}{-A_{d,x}} \\ 0 & 0 & 0 & \rho_x \end{pmatrix}$$

$$\tilde{S} = \begin{pmatrix} 1 & 0 & 0 & \frac{Q_{1,1}}{-A_{d,x}} s_{\sigma,x} \\ \frac{Q_{2,2}}{Q_{1,1}} s_{q,\sigma} & 1 & 0 & \frac{Q_{2,2}}{-A_{d,x}} s_{q,x} \\ \frac{1}{Q_{1,1}} s_{d,\sigma} & \frac{1}{Q_{2,2}} s_{d,q} & 1 & \frac{1}{-A_{d,x}} s_{d,x} \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \tilde{\epsilon}_{t+1} = \begin{pmatrix} Q_{1,1} \\ Q_{2,2} \\ 1 \\ -A_{d,x} \end{pmatrix} \odot \epsilon_{t+1}.$$

B.4.1 Separate Volatility Dynamics

In the alternative setup with separate volatility dynamic, ρ, ϵ_{t+1} and S may be expressed as,

$$\rho = \begin{pmatrix} \rho_\sigma & 0 & 0 & 0 \\ 0 & \rho_q & 0 & 0 \\ 0 & 0 & \rho_d & \phi_{dx} \\ 0 & 0 & 0 & \rho_x \end{pmatrix} \qquad \epsilon_{t+1} = \begin{pmatrix} \varphi_\sigma \sigma_t z_{\sigma,t+1} \\ \varphi_q \sqrt{q_t} z_{q,t+1} \\ \varphi_d \sigma_t z_{d,t+1} \\ \sqrt{q_t} z_{x,t+1} \end{pmatrix} \qquad S = \begin{pmatrix} 1 & 0 & 0 & s_{\sigma,x} \\ 0 & 1 & 0 & s_{q,x} \\ s_{d,\sigma} & s_{d,q} & 1 & s_{d,x} \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{B.52}$$

$$X_t = \mu_X + Q f_t \qquad Q = \begin{pmatrix} Q_{1,1} & Q_{1,2} & 0 & 0 \\ Q_{2,1} & Q_{2,2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -A_{d,\sigma} & -A_{d,q} & -A_{d,d} & -A_{d,x} \end{pmatrix}$$

Consequently,,

$$Q^{-1} = \begin{pmatrix} \frac{Q_{2,2}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & \frac{-Q_{1,2}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & 0 & 0 \\ \frac{-Q_{2,1}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & \frac{Q_{1,1}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{Q_{2,1}A_{d,q}-Q_{2,2}A_{d,\sigma}}{(Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1})A_{d,x}} & \frac{-Q_{1,1}A_{d,q}+Q_{1,2}A_{d,\sigma}}{(Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1})A_{d,x}} & \frac{-A_{d,d}}{A_{d,x}} & -\frac{1}{A_{d,x}} \end{pmatrix}$$

$$\frac{1}{\text{diag}(Q^{-1})} = \begin{pmatrix} \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{Q_{2,2}} \\ \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{Q_{1,1}} \\ 1 \\ -A_{d,x} \end{pmatrix}$$

Combining these expressions, it follows that

$$B = \begin{pmatrix} 1 & \frac{-Q_{1,2}}{Q_{2,2}} & 0 & 0 \\ \frac{-Q_{2,1}}{Q_{1,1}} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{-Q_{2,1}A_{d,q}+Q_{2,2}A_{d,\sigma}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & \frac{+Q_{1,1}A_{d,q}-Q_{1,2}A_{d,\sigma}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & A_{d,d} & 1 \end{pmatrix} \qquad \tilde{\rho} = \begin{pmatrix} \rho_\sigma & 0 & 0 & 0 \\ 0 & \rho_q & 0 & 0 \\ 0 & 0 & \rho_d & \frac{\phi_{dx}}{-A_{d,x}} \\ 0 & 0 & 0 & \rho_x \end{pmatrix}$$

$$\tilde{S} = \begin{pmatrix} 1 & 0 & 0 & \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{-Q_{2,2}A_{d,x}}s_{\sigma,x} \\ 0 & 1 & 0 & \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{-Q_{1,1}A_{d,x}}s_{q,x} \\ \frac{Q_{2,2}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}s_{d,\sigma} & \frac{Q_{1,1}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}s_{d,q} & 1 & \frac{1}{-A_{d,x}}s_{d,x} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\tilde{\epsilon}_{t+1} = \begin{pmatrix} \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{Q_{2,2}} \\ \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{Q_{1,1}} \\ 1 \\ -A_{d,x} \end{pmatrix} \odot \epsilon_{t+1}.$$

B.4.2 Stochastic Volatility in the Long-Run

In the alternative setup with stochastic volatility in the long-run drift, ρ, ϵ_{t+1} and S may be expressed as,

$$\rho = \begin{pmatrix} \rho_\sigma & 1 & 0 & 0 \\ 0 & \rho_q & 0 & 0 \\ 0 & 0 & \rho_d & \phi_{dx} \\ 0 & 0 & 0 & \rho_x \end{pmatrix} \qquad \epsilon_{t+1} = \begin{pmatrix} \varphi_\sigma \sigma_t z_{\sigma,t+1} \\ \varphi_q \sqrt{q_t} z_{q,t+1} \\ \varphi_d \sigma_t z_{d,t+1} \\ \varphi_x \sigma_t z_{x,t+1} \end{pmatrix} \qquad S = \begin{pmatrix} 1 & 0 & 0 & s_{\sigma,x} \\ 0 & 1 & 0 & s_{q,x} \\ s_{d,\sigma} & s_{d,q} & 1 & s_{d,x} \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{B.53}$$

$$X_t = \mu_X + Q f_t \qquad Q = \begin{pmatrix} Q_{1,1} & Q_{1,2} & 0 & 0 \\ Q_{2,1} & Q_{2,2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -A_{d,\sigma} & -A_{d,q} & -A_{d,d} & -A_{d,x} \end{pmatrix}$$

Consequently,

$$Q^{-1} = \begin{pmatrix} \frac{Q_{2,2}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & \frac{-Q_{1,2}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & 0 & 0 \\ \frac{-Q_{2,1}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & \frac{Q_{1,1}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{Q_{2,1}A_{d,q}-Q_{2,2}A_{d,\sigma}}{(Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1})A_{d,x}} & \frac{-Q_{1,1}A_{d,q}+Q_{1,2}A_{d,\sigma}}{(Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1})A_{d,x}} & \frac{-A_{d,d}}{A_{d,x}} & -\frac{1}{A_{d,x}} \end{pmatrix}$$

$$\frac{1}{\text{diag}(Q^{-1})} = \begin{pmatrix} \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{Q_{2,2}} \\ \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{Q_{1,1}} \\ 1 \\ -A_{d,x} \end{pmatrix}$$

Combining these expressions, it follows that

$$B = \begin{pmatrix} 1 & \frac{-Q_{1,2}}{Q_{2,2}} & 0 & 0 \\ \frac{-Q_{2,1}}{Q_{1,1}} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{-Q_{2,1}A_{d,q}+Q_{2,2}A_{d,\sigma}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & \frac{+Q_{1,1}A_{d,q}-Q_{1,2}A_{d,\sigma}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} & A_{d,d} & 1 \end{pmatrix} \qquad \tilde{\rho} = \begin{pmatrix} \rho_\sigma & \frac{Q_{1,1}}{Q_{2,2}} & 0 & 0 \\ 0 & \rho_q & 0 & 0 \\ 0 & 0 & \rho_d & \frac{\phi_{dx}}{-A_{d,x}} \\ 0 & 0 & 0 & \rho_x \end{pmatrix}$$

$$\tilde{S} = \begin{pmatrix} 1 & 0 & 0 & \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{-Q_{2,2}A_{d,x}} s_{\sigma,x} \\ 0 & 1 & 0 & \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{-Q_{1,1}A_{d,x}} s_{q,x} \\ \frac{Q_{2,2}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} s_{d,\sigma} & \frac{Q_{1,1}}{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}} s_{d,q} & 1 & \frac{1}{-A_{d,x}} s_{d,x} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\tilde{\epsilon}_{t+1} = \begin{pmatrix} \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{Q_{2,2}} \\ \frac{Q_{1,1}Q_{2,2}-Q_{1,2}Q_{2,1}}{Q_{1,1}} \\ 1 \\ -A_{d,x} \end{pmatrix} \odot \epsilon_{t+1}.$$

Table B.1: Structural Factor GARCH Estimates—Separate Volatility Dynamics

The table reports the "structural" factor GARCH estimates for the alternative setup with separate volatility dynamics described in Sections B.3.1 and B.4.1, with the three restrictions: $A_{d,d} = \frac{\rho_d}{1-\kappa_{d,1}\rho_d}$, $\Gamma_{4,4} + \Upsilon_{4,4} = \rho_q$, and $\Gamma_{3,3} = 0$. The resulting J-test with 7 degrees-of-freedom for the GMM-based estimation equals 26.31, corresponding to a p-value 0.0004.

B			ERV_{t+1}		VRP_{t+1}		Δd_{t+1}		d_{t+1}/p_{t+1}	
ERV_{t+1}			1		-0.490	(0.117)	0		0	
VRP_{t+1}			-0.022	(0.030)	1		0		0	
Δd_{t+1}			0		0		1		0	
d_{t+1}/p_{t+1}			-0.110	(0.141)	-1.595	(0.063)	-0.158		1	
$\tilde{\rho}$	constant		ERV_t		VRP_t		Δd_t		d_t/p_t	
ERV_{t+1}	0.009	(0.003)	0.827	(0.089)	0		0		0	
VRP_{t+1}	0.008	(0.002)	0		0.312	(0.071)	0		0	
Δd_{t+1}	-0.002	(0.016)	0		0		-0.187	(0.035)	-0.001	(0.004)
d_{t+1}/p_{t+1}	-0.080	(0.029)	0		0		0		0.980	(0.008)
$\tilde{S}$			$\tilde{\epsilon}_{\sigma^2_{t+1}}$		$\tilde{\epsilon}_{q_{t+1}}$		$\tilde{\epsilon}_{\Delta d_{t+1}}$		$\tilde{\epsilon}_{x,t+1}$	
ERV_{t+1}			1		0		0		0.316	(0.038)
VRP_{t+1}			0		1		0		-0.245	(0.017)
Δd_{t+1}			-0.387	(0.080)	-0.134	(0.160)	1		0.095	(0.034)
d_{t+1}/p_{t+1}			0		0		0		1	
$\tilde{\epsilon}$	ϖ_u		Γ		Υ					
ERV_{t+1}	0.001	(0.000)	0.153	(0.360)	0.388	(0.138)				
VRP_{t+1}	0.000	(0.000)	0.537	(0.335)	0.116	(0.079)				
Δd_{t+1}	0.001	(0.000)	0		0.449	(0.115)				
d_{t+1}/p_{t+1}	0.002	(0.000)	0.167		0.144	(0.106)				

Table B.2: Structural Model Implications—Separate Volatility Dynamics

The table reports the contemporaneous matrix Φ_0, the reduced form matrix Φ, and the return equation, implied by the alternative "structural" factor GARCH model defined in Sections B.3.1 and B.4.1).

$\Phi_0^{-1} \equiv B^{-1}\tilde{S}$			ERV_{t+1}		VRP_{t+1}		Δd_{t+1}		d_{t+1}/p_{t+1}	
ERV_{t+1}			1.011	(0.015)	0.496	(0.118)	0		0.198	(0.049)
VRP_{t+1}			0.022	(0.030)	1		0		-0.241	(0.017)
Δd_{t+1}			-0.387	(0.080)	-0.134	(0.160)	1		0.095	(0.034)
d_{t+1}/p_{t+1}			0.085	(0.134)	1.646	(0.079)	0.158	(0.025)	0.652	(0.047)
$\Phi \equiv B^{-1}\tilde{\rho}B$	constant		ERV_t		VRP_t		Δd_t		d_t/p_t	
ERV_{t+1}	0.013	(0.003)	0.833	(0.091)	-0.255	(0.075)	0		0	
VRP_{t+1}	0.008	(0.001)	0.012	(0.016)	0.306	(0.070)	0.000	(0.000)	0.000	(0.000)
Δd_{t+1}	-0.002	(0.016)	0.000	(0.000)	0.002	(0.006)	-0.187	(0.035)	-0.001	(0.004)
d_{t+1}/p_{t+1}	-0.066	(0.030)	0.002	(0.045)	-1.103	(0.130)	-0.185	(0.035)	0.980	(0.008)
	GMM Implied Return Equation									
$r_{t,t+1}$	constant		ERV_t		VRP_t		Δd_t		d_t/p_t	
	0.062	(0.029)	-0.002	(0.044)	1.074	(0.127)	-0.007	(0.002)	0.046	(0.009)
			$\tilde{\epsilon}_{\sigma^2_{t+1}}$		$\tilde{\epsilon}_{q_{t+1}}$		$\tilde{\epsilon}_{\Delta d_{t+1}}$		$\tilde{\epsilon}_{x,t+1}$	
stru-shocks			-0.470	(0.213)	-1.734	(0.178)	0.846	(0.041)	-0.539	(0.036)

Table B.3: Structural Factor GARCH Estimates—Long-Run Stochastic Volatility

The table reports the "structural" factor GARCH estimates for the alternative setup with long-run stochastic volatility described in Sections B.3.2 and B.4.2, with the two restrictions: $A_{d,d} = \frac{\rho_d}{1-\kappa_{d,1}\rho_d}$ and $\Gamma_{3,3} = 0$. The resulting J-test with 6 degrees-of-freedom for the GMM-based estimation equals 37.02, corresponding to a p-value 0.0000.

B			ERV_{t+1}		VRP_{t+1}		Δd_{t+1}		d_{t+1}/p_{t+1}	
ERV_{t+1}			1		0.000	(0.189)	0		0	
VRP_{t+1}			0.120	(0.039)	1		0		0	
Δd_{t+1}			0		0		1		0	
d_{t+1}/p_{t+1}			-0.016	(0.086)	-2.007	(0.156)	-0.249		1	
$\tilde{\rho}$	constant		ERV_t		VRP_t		Δd_t		d_t/p_t	
ERV_{t+1}	0.003	(0.002)	1.001	(0.077)	-0.070	(0.255)	0		0	
VRP_{t+1}	0.006	(0.001)	0		0.609	(0.079)	0		0	
Δd_{t+1}	-0.000	(0.014)	0		0		-0.329	(0.040)	-0.001	(0.004)
d_{t+1}/p_{t+1}	-0.075	(0.028)	0		0		0		0.982	(0.007)
$\tilde{S}$			$\tilde{\epsilon}_{\sigma^2_{t+1}}$		$\tilde{\epsilon}_{q_{t+1}}$		$\tilde{\epsilon}_{\Delta d_{t+1}}$		$\tilde{\epsilon}_{x,t+1}$	
ERV_{t+1}			1		0		0		0.332	(0.036)
VRP_{t+1}			0		1		0		-0.186	(0.016)
Δd_{t+1}			-0.524	(0.061)	-0.069	(0.147)	1		0.061	(0.024)
d_{t+1}/p_{t+1}			0		0		0		1	
$\tilde{\epsilon}$	ϖ_u		Γ		Υ					
ERV_{t+1}	0.001	(0.000)	0.001	(0.051)	0.776	(0.062)				
VRP_{t+1}	0.000	(0.000)	0.000	(0.139)	0.322	(0.147)				
Δd_{t+1}	0.001	(0.000)	0		0.454	(0.095)				
d_{t+1}/p_{t+1}	0.002	(0.000)	0.766	(0.061)	0.160	(0.041)				

Table B.4: Structural Model Implications–Long-Run Stochastic Volatility

The table reports the contemporaneous matrix Φ_0, the reduced form matrix Φ, and the return equation, implied by the alternative "structural" factor GARCH model in Sections B.3.2 and B.4.2.

$\Phi_0^{-1} \equiv B^{-1}\tilde{S}$			ERV_{t+1}		VRP_{t+1}		Δd_{t+1}		d_{t+1}/p_{t+1}	
ERV_{t+1}			1.000	(0.023)	-0.000	(0.189)	0		0.332	(0.032)
VRP_{t+1}			-0.120	(0.037)	1		0		-0.226	(0.025)
Δd_{t+1}			-0.524	(0.061)	-0.069	(0.147)	1		0.061	(0.024)
d_{t+1}/p_{t+1}			-0.356	(0.052)	1.990	(0.158)	0.249	(0.023)	0.568	(0.051)
$\Phi \equiv B^{-1}\tilde{\rho}B$	constant		ERV_t		VRP_t		Δd_t		d_t/p_t	
ERV_{t+1}	0.003	(0.002)	0.993	(0.089)	-0.070	(0.194)	0		0	
VRP_{t+1}	0.005	(0.001)	-0.046	(0.020)	0.618	(0.069)	0		0	
Δd_{t+1}	-0.000	(0.014)	0.000	(0.000)	0.003	(0.007)	-0.328	(0.040)	-0.001	(0.004)
d_{t+1}/p_{t+1}	-0.065	(0.029)	-0.093	(0.041)	-0.732	(0.174)	-0.326	(0.041)	0.982	(0.007)
	GMM Implied Return Equation									
$r_{t,t+1}$	constant		ERV_t		VRP_t		Δd_t		d_t/p_t	
	0.063	(0.027)	0.090	(0.040)	0.714	(0.169)	-0.011	(0.002)	0.045	(0.008)
			$\tilde{\epsilon}_{\sigma^2_{t+1}}$		$\tilde{\epsilon}_{q_{t+1}}$		$\tilde{\epsilon}_{\Delta d_{t+1}}$		$\tilde{\epsilon}_{x,t+1}$	
stru-shocks			-0.178	(0.124)	-2.004	(0.181)	0.758	(0.050)	-0.491	(0.035)

Appendix C

Supplementary Appendix to "Tail Risk and Equity Risk Premia"

C.1 Model Solution

In this section, I explain in detail how to derive the closed-form solutions for the pricing kernel and the equity risk premia. The overall solution method is typical as in Bansal and Yaron (2004b) and Drechsler and Yaron (2011b), among others. The solutions derived here are crucial to understanding the risk factors and the way in which they are priced.

C.1.1 Equity Risk Premium

I start by conjecturing the hedging demand as a linear function of the jump intensity, the volatility factor and the risk free rate,

$$h_t = A_0 + A_q q_t + A_{\lambda,t}\lambda_t + A_f r_{f,t}. \tag{C.1}$$

where the risk free rate follows a CIR model $dr_{f,t} = \kappa_f(\mu_f - r_{f,t})dt + \varphi_f\sqrt{r_{f,t}}dW_{f,t}$, $W_{f,t}$ is an independent Brownian motion of $W_{i,t}, W_{c^\perp,t}$, $W_{m^\perp,t}$ and $W_{q,t}$, and the

loading on the intensity is time-varying $\lambda_t dA_{\lambda,t} = \mu_{A_\lambda,t}\lambda_t + \varphi_{A_\lambda} dW_{A_{\lambda,t},t}$. In turn, the stochastic discount factor M_t follows a jump diffusion model as well,

$$dlnM_t = -r_{f,t}dt - m_t dt + \sigma_t dW_t + \int_{\mathbb{B}} (-\gamma_{J,t}x + \phi_{\lambda,t}) J(dt, dx), \tag{C.2}$$

$$\sigma_t dW_t = -\gamma\sigma_c dW_{c^\perp,t} + \sqrt{q_t}\left(\gamma\varphi_c + \phi_q\right) dW_{q,t} + \phi_f\sqrt{r_{f_t}} dW_{f,t} + \phi_{A_\lambda} dW_{A_{\lambda,t},t},$$

$$m_t = 0.5\left(\phi_{A_\lambda}^2 + \gamma^2\sigma_c^2 + (\gamma\varphi_c + \phi_q)^2 q_t + \phi_f^2 r_{f,t}\right) + \lambda_t \int_{\mathbb{B}} (e^{-\gamma_{J,t}x + \phi_{\lambda,t}} - 1) f_t(x) dx.$$

where $\phi_q = \frac{\theta}{\Psi} A_q \varphi_q$, $\phi_{\lambda,t} = \frac{\theta}{\Psi} A_{\lambda,t}\varphi_\lambda$, $\phi_f = \frac{\theta}{\Psi} A_f \varphi_f$ and $\phi_{A_\lambda} = \frac{\theta}{\Psi}\varphi_{A_\lambda}$ are pricing coefficients. The drift term m_t ensures the no arbitrage condition holds for bond pricing, and W_t summarizes the Brownian motions from hedging demand h_t and the total wealth return $R_{c,t}$. Taken together, based on a no-arbitrage condition $d(R_{c,t}M_t)$=0, the solution for the drift term $a_{c,t}$ has the following form,

$$\begin{aligned} a_{c,t} &= r_{f,t} + \gamma\sigma_c^2 + \varphi_c(\gamma\varphi_c + \phi_q) q_t \\ &\quad + \lambda_t \left(\int_B (e^{\frac{\gamma_{J,t}}{\gamma}x} - 1) f_t(x) dx - \int_B (e^{\frac{\gamma_{J,t}}{\gamma}x} - 1) e^{-\gamma_{J,t}x + \phi_{\lambda,t}} f_t(x) dx \right). \end{aligned} \tag{C.3}$$

Lastly, to solve the pricing coefficients, I use equations (C.1)-(C.3) to substitute out corresponding terms in the Euler equation (3.7), which must hold regardless of the realization of the state vector $[q_t,\, r_{f,t},\, \lambda_t]$. Therefore, it follows that,

$$0 = (\kappa_q + \frac{1-k_1}{k_1}) A_q + (\Psi/\theta)\left(-0.5\phi_q^2 + \left(\gamma\varphi_c^2(\gamma - 0.5) - 0.5\gamma^2\varphi_c^2\right)\right),$$

$$\begin{aligned} 0 &= (\kappa_\lambda + \frac{1-k_1}{k_1}) A_{\lambda,t} + (\Psi/\theta)\left(\gamma \int_{\mathbb{R}} (1 - e^{\frac{\gamma_{J,t}}{\gamma}x}) e^{-\gamma_{J,t}x + \phi_{\lambda,t}} f^{\mathbb{P}}(x) dx \right. \\ &\quad \left. - \int_{\mathbb{R}} (1 - e^{\gamma_{J,t}x - \phi_{\lambda,t}}) e^{-\gamma_{J,t}x + \phi_{\lambda,t}} f^{\mathbb{P}}(x) dx, \right. \end{aligned}$$

$$0 = (\kappa_f + \frac{1-k_1}{k_1}) A_f + (\Psi/\theta)\left(\gamma - 0.5\phi_f^2 - 1\right).$$

Or equivalently, the pricing coefficients $\phi_{\cdot}$ have the following expressions,

$$\phi_q = (\kappa_q + \frac{1-k_1}{k_1})/\varphi_q - \left((\kappa_q + \frac{1-k_1}{k_1})^2/\varphi_q^2 + \gamma\varphi_c^2(\gamma-1)\right)^{1/2},$$

$$\phi_{\lambda,t} = \frac{\mu_{A_\lambda,t}}{\kappa_\lambda + \frac{1-k_1}{k_1}} - \varphi_\lambda \frac{\gamma\int_{\mathbb{B}}(1-e^{\frac{\gamma_{J,t}}{\gamma}x})e^{-\gamma_{J,t}x+\phi_{\lambda,t}}f_t(x)dx - \int_{\mathbb{B}}(1-e^{\gamma_{J,t}x-\phi_{\lambda,t}})e^{-\gamma_{J,t}x+\phi_{\lambda,t}}f_t(x)dx}{\kappa_\lambda + \frac{1-k_1}{k_1}},$$

$$\phi_f = (\kappa_f + \frac{1-k_1}{k_1})/\varphi_f - \left((\kappa_f + \frac{1-k_1}{k_1})^2/\varphi_f^2 + 2(\gamma-1)\right)^{1/2}.$$

There are several things worth noticing: first, the sign of all pricing coefficients are related to the magnitude of risk aversion γ, but not the elasticity parameter Ψ. Second, if $\gamma > 1$, then $\phi_f < 0$, which means the risk free rate serves as a good hedging portfolio and is negatively priced. Third, despite $\gamma > 1$ leading to a negative ϕ_q, volatility is always positively priced as $\gamma\varphi_c + \phi_q > 0$. Lastly , the intensity premium $\phi_{\lambda,t}$ is time varying, due to its stochastic long-run mean $\mu_{A_\lambda,t}$ and the shifting jump shape.

C.1.2 Variance Risk Premia

Variance risk premium $VRP_{t,T}$ is the difference between the risk neutral and the physical expectations of the quadratic variance, and is also naturally decomposed into two parts $VRP_{t,T} = VRP^{cv}_{t,T} + VRP^{J}_{t,T}$,

$$VRP^{J}_{t,T} \equiv -E_t\int_t^T \lambda_s \int_B x^2 f_s(x)dxds + E_t^{\mathbb{Q}}\int_t^T \lambda_s^{\mathbb{Q}}\int_B x^2 f_s^{\mathbb{Q}}(x)dxds, \quad \text{(C.4)}$$

$$VRP^{cv}_{t,T} \equiv -E_t\int_t^T \left(\sigma_m^2 + \varphi_m^2 q_s\right)ds + E_t^{\mathbb{Q}}\int_t^T \left(\sigma_m^2 + \varphi_m^2 q_s\right)ds, \quad \text{(C.5)}$$

$$= -E_t\int_t^T \varphi_m^2 q_s ds + E_t^{\mathbb{Q}}\int_t^T \varphi_m^2 q_s ds. \quad \text{(C.6)}$$

Next, to go one step further, the conditional mean of the square root process q_t has the following expression,

$$E_t q_{t+s} = e^{-\kappa_q s}(q_t - \mu_q) + \mu_q. \tag{C.7}$$

and the difference between $\mathbb{Q}$ and $\mathbb{P}$ expectation is,

$$E_t^{\mathbb{Q}} q_{t+s} - E_t q_{t+s} = q_t(e^{-\kappa_q^{\mathbb{Q}} s} - e^{-\kappa_q s}) + e^{-\kappa_q^{\mathbb{Q}} s}\mu_q^{\mathbb{Q}} - e^{-\kappa_q s}\mu_q + \mu_q^{\mathbb{Q}} - \mu_q. \tag{C.8}$$

where $\kappa_q^{\mathbb{Q}} = \kappa_q - \varphi_q(\gamma\varphi_c + \phi_q)$, $\quad \kappa_q^{\mathbb{Q}}\mu_q^{\mathbb{Q}} = \kappa_q\mu_q$. As such, both the expected future return variance $E_t CV$ and the continuous part of the variance risk premium VRP_t^{cv} are affine in latent factor q_t,

$$E_t^{\mathbb{P}} CV \equiv E_t^{\mathbb{P}} \int_t^T \sigma_c^2 + \sigma_m^2 + \varphi_m^2 q_s ds = a_{4,0} + a_{4,3} q_t, \tag{C.9}$$

$$VRP_t^{cv} = a_{3,0} + a_{3,3} q_t. \tag{C.10}$$

where $ey(k) = \frac{e^{-k(T-t)}-1}{-k}$ and,

$a_{4,0} = \sigma_c^2(T-t) + \sigma_m^2(T-t) + \varphi_m^2(T - t - ey(\kappa_q)), \quad a_{4,3} = ey(\kappa_q),$

$a_{3,0} = \varphi_m^2[(T-t)(\mu_q^{\mathbb{Q}} - \mu_q) - ey(\kappa_q^{\mathbb{Q}})\mu_q^{\mathbb{Q}} + ey(\kappa_q)\mu_q], \qquad a_{3,3} = \varphi_m^2[ey(\kappa_q^{\mathbb{Q}}) - ey(\kappa_q)].$

Lastly, I turn to the jump part of the variance risk premium $VRP_{t,T}^J$. By assuming the intensity premium and the shape parameters are unchanged within a short period of time $T - t$, the expectations of the self-exciting jump intensity are,

$$E_t \lambda_{t+s} = e^{-\tilde{\kappa}_\lambda s}(\lambda_t - \tilde{\mu}_\lambda) + \tilde{\mu}_\lambda, \tag{C.11}$$

$$E_t^{\mathbb{Q}} \lambda_{t+s} = e^{-\tilde{\kappa}_\lambda^{\mathbb{Q}} s}(\lambda_t - \tilde{\mu}_\lambda^{\mathbb{Q}}) + \tilde{\mu}_\lambda^{\mathbb{Q}}. \tag{C.12}$$

where $\tilde{\kappa}_\lambda = \kappa_\lambda - \varphi_\lambda$, $\tilde{\kappa}_\lambda \tilde{\mu}_\lambda = \kappa_\lambda \mu_\lambda$ and $\tilde{\kappa}^{\mathbb{Q}}_\lambda = \kappa_\lambda - \varphi_\lambda e^{\phi_{\lambda,t}} \int_B e^{-\gamma_{J,t}x} f_t(x)dx$, $\tilde{\kappa}^{\mathbb{Q}}_\lambda \tilde{\mu}^{\mathbb{Q}}_\lambda = \kappa_\lambda \mu_\lambda$. By integrating over this time period,

$$
\begin{aligned}
E_t \int_t^T \lambda_s ds &= a_{1,0} + a_{1,1}\lambda_t, && \text{(C.13)} \\
E^{\mathbb{Q}}_t \int_t^T \lambda^{\mathbb{Q}}_s ds &= e^{\phi_{\lambda,t}} \int_B e^{-\gamma_{J,t}x} f_t(x)dx \times E^{\mathbb{Q}}_t \int_t^T \lambda_s ds \\
&= a_{2,0} e^{\phi_{\lambda,t}} \int_B e^{-\gamma_{J,t}x} f_t(x)dx + a_{2,2}\lambda^{\mathbb{Q}}_t. && \text{(C.14)}
\end{aligned}
$$

where $a_{1,0} = T - t - ey(\tilde{\kappa}_\lambda)$, $a_{1,1} = ey(\tilde{\kappa}_\lambda)$ and $a_{2,0} = T - t - ey(\tilde{\kappa}^{\mathbb{Q}}_\lambda)$, $a_{2,2} = ey(\tilde{\kappa}^{\mathbb{Q}}_\lambda)$. The $\mathbb{P}$- measure expectation in equation (C.4) becomes $E_t \int_t^T \lambda_s ds \times \int_{\mathbb{B}} x^2 f_t(x)dx = (a_{1,0} + a_{1,1}\lambda_t) \times \int_{\mathbb{B}} x^2 f_t(x)dx$, and the $\mathbb{Q}$- measure expectation becomes $E^{\mathbb{Q}}_t \int_t^T \lambda^{\mathbb{Q}}_s ds \times \int_{\mathbb{B}} x^2 f^{\mathbb{Q}}_t(x)dx = (a_{2,0} e^{\phi_{\lambda,t}} \int_{\mathbb{B}} e^{-\gamma_{J,t}x} f_t(x)dx + a_{2,2}\lambda^{\mathbb{Q}}_t) \times \int_{\mathbb{B}} x^2 f^{\mathbb{Q}}_t(x)dx$. For simplicity, further assume the intensity process is very persistent under both measures, $\frac{VRP^J_{t,T}}{T-t} \approx \lambda^{\mathbb{Q}}_t \times \int_{\mathbb{B}} x^2 f^{\mathbb{Q}}_t(x)dx - \lambda_t \times \int_{\mathbb{B}} x^2 f_t(x)dx$.

The implied variance IV_t quantified by VIX (CBOE), contains a bias from the jump process.[1] It measures the expected difference between the gross return and the log gross return under the risk neutral measure,

$$
\begin{aligned}
IV_{t,T} &\equiv 2E^{\mathbb{Q}}_t[\int_t^T \frac{P_T}{P_t} - log\frac{P_T}{P_t}] \\
&= E^{\mathbb{Q}}_t[CV_{t,T}] + 2E^{\mathbb{Q}}_t \int_t^T \int_{\mathbb{B}} (e^x - x - 1) v^{\mathbb{Q}}_s(dx)ds. && \text{(C.15)}
\end{aligned}
$$

[1] Implied variance is commonly used in empirical asset pricing studies, see e.g. Bollerslev et al. (2009), Du and Kapadia (2012), among others.

where P_t is the price for the aggregate market. Therefore,

$$\widehat{E_t^{\mathbb{Q}}[CV_{t,T}]} = IV_{t,T} - 2a_J\widehat{\lambda_t^{\mathbb{Q}}} \times \int_{\mathbb{B}} (e^x - x - 1)\widehat{f_t^{\mathbb{Q}}(x)}dx, \tag{C.16}$$

$$\widehat{VRP_{t,T}^{cv}} = \widehat{E_t^{\mathbb{Q}}[CV_{t,T}]} - \widehat{E_t[CV_{t,T}]}. \tag{C.17}$$

where a_J is approximately $T-t$, $\widehat{E_t[CV_{t,T}]}$ is estimated together with EJX in equation (3.29) by using high frequency data in a HAR-VAR Kalman filter approach (Appendix C). To construct the volatility part of the equity risk premium, a parametric calibration is needed besides $\widehat{VRP_{t,T}^{cv}}$.

C.1.3 Implied Consumption

Based on this setup, the implied log consumption takes the following expression,

$$\begin{aligned} dlnC_t =& -\Psi\delta dt - \frac{\Psi}{\theta}\left(dlnM_t + (\theta-1)lnR_{c,t}\right) \\ =& \Big(-\Psi\delta + \frac{\Psi}{\theta}\{m_t + r_{f,t} + (\theta-1)[a_{c,t} - 0.5(\sigma_c^2 + \varphi_c^2 q_t) \\ & - \int_{\mathbb{B}} (e^{\frac{\gamma_{J,t}}{\gamma}x} - 1)J(dt,dx)]\}\Big)dt + \int_{\mathbb{B}} \left[\frac{\gamma_{J,t}}{\gamma}x - \phi_{\lambda,t}\frac{\Psi}{\theta}\right] J(dt,dx) \\ & + \sigma_c dW_{c^\perp,t} - (\varphi_c + \frac{\Psi}{\theta}\phi_q)\sqrt{q_t}dW_{q,t} - \frac{\Psi}{\theta}\phi_f\sqrt{r_{f_t}}dW_{f,t} - \varphi_{A_\lambda}dW_{A_{\lambda,t},t}. \end{aligned} \tag{C.18}$$

The shock structure shows how consumption is co-moving with different sources of risk. The first shock component represents the endowment risk, which is a one-to-one mapping with the total wealth return. The next component says that, if Ψ(IES) is larger than 1 ($\theta < 0$), then when volatility uncertainty increases, consumption growth decreases (since $\phi_q < 0$). If Ψ(IES) is smaller than 1 ($\theta > 0$), to ensure the same relation, the parameter space for Ψ should be more restricted, $1 > \Psi > 1 - \varphi_c(1-\gamma)/\phi_q$.

The third component captures the large discontinuous movement in the economy.

Since the intensity premium is positive $\phi_{\lambda,t} > 0$,[2] it requires $\Psi > 1$ to ensure the consumption is less affected by left-jumps than the total wealth return. When $\Psi > 1$, the fourth component suggests that a positive monetary policy shock coincides with a decrease in consumption growth ($\phi_f < 0$). Lastly, regardless of the choice of Ψ(IES), a positive shock to the intensity risk premium ($\phi_{\lambda,t} > 0$) slows down the consumption growth rate.

Note that since $0.5\frac{\Psi}{\theta}\gamma^2 + \frac{\Psi}{\theta}(\theta-1)(\gamma-0.5) + 0.5 = 0.5\gamma(\Psi+1)$, by Ito's lemma, the geometric consumption growth can be expressed as follows,

$$\begin{aligned}
\frac{dC_t}{C_{t^-}} &= [-\Psi\delta + \left(0.5\frac{\Psi}{\theta}(\frac{\Psi}{\theta}+1)\phi_f^2 + \Psi\right) r_{f,t} + 0.5\frac{\Psi}{\theta}(\frac{\Psi}{\theta}+1)\phi_{A_\lambda}^2]dt \\
&\quad + [0.5\gamma(1+\Psi)\sigma_c^2 + \{0.5\frac{\Psi}{\theta}(\frac{\Psi}{\theta}+1)\phi_q^2 + 0.5\gamma(1+\Psi)\varphi_c^2 + (\frac{\Psi}{\theta}+1)\varphi_c\phi_q\}q_t]dt \\
&\quad - (1-\frac{\Psi\gamma}{\theta})\int_{\mathbb{B}}(e^{\frac{\gamma_{J,t}}{\gamma}x}-1)v_t^{\mathbb{Q}}(x)dx + \frac{\Psi}{\theta}\int_{\mathbb{B}}(1-e^{\gamma\frac{\gamma_{J,t}}{\gamma}x-\phi_{\lambda,t}})v_t^{\mathbb{Q}}(x)dx \\
&\quad + \int_{\mathbb{B}}(e^{\frac{\gamma_{J,t}}{\gamma}x-A_{\lambda,t}\varphi_\lambda}-1)J(dt,dx) + \sigma_c dW_{c^\perp,t} - [\varphi_c + \frac{\Psi}{\theta}\phi_q]\sqrt{q_t}dW_{q,t} \\
&\quad - \frac{\Psi}{\theta}\phi_f\sqrt{r_{f_t}}dW_{f,t} - \varphi_{A_\lambda}dW_{A_{\lambda,t},t}. \qquad \text{(C.19)}
\end{aligned}$$

[2] This is based on the estimation of $\phi_{\lambda,t}$ in section 5.

with implied unconditional mean,

$$
\begin{aligned}
E\int_t^{t+\Delta T}\frac{dC_t}{C_{t^-}} = &-\Delta T\Psi\delta + E\int_t^{t+\Delta T}\left(0.5\frac{\Psi}{\theta}(\frac{\Psi}{\theta}+1)\phi_f^2+\Psi\right)r_{f,s}ds \\
&+ 0.5\frac{\Psi}{\theta}(\frac{\Psi}{\theta}+1)\phi_{A_\lambda}^2\Delta T + 0.5\gamma(1+\Psi)\sigma_c^2\Delta T \\
&+ \left(0.5\frac{\Psi}{\theta}(\frac{\Psi}{\theta}+1)\phi_q^2 + 0.5\gamma(1+\Psi)\varphi_c^2 + (\frac{\Psi}{\theta}+1)\varphi_c\phi_q\right)E\int_t^{t+\Delta T}q_s ds \\
&+ E\int_t^{t+\Delta T}[\int_{\mathbb{B}}(e^{\frac{\gamma_{J,t}}{\gamma}x - A_{\lambda,t}\varphi_\lambda}-1)v_s(x)dx \\
&- (1-\frac{\Psi\gamma}{\theta})\int_{\mathbb{B}}(e^{\frac{\gamma_{J,t}}{\gamma}x}-1)v_s^{\mathbb{Q}}(x)dx + \frac{\Psi}{\theta}\int_{\mathbb{B}}(1-e^{\gamma_{J,t}x-\phi_{\lambda,t}})v_s^{\mathbb{Q}}(x)dx].
\end{aligned}
$$

and a quadratic variance,

$$
\begin{aligned}
QV_{t,T} = &\int_t^T\int_{\mathbb{B}}\left(\frac{\gamma_{J,t}}{\gamma}x - A_{\lambda,t}\varphi_\lambda\right)^2 v_t(x)dx + \sigma_c^2\Delta T + \int_t^T(\varphi_c+\frac{\Psi}{\theta}\phi_q)^2 q_s ds \\
&+ \int_t^T\frac{\Psi^2}{\theta^2}\phi_f^2 r_{f,s}ds + \frac{\Psi^2}{\theta^2}\phi_{A_\lambda}^2\Delta T.
\end{aligned}
$$

C.2 Option Data Cleaning Rule

The cleaning filter for options are in line with the option literature. I first drop those observations with non-positive bid, and those with non-decreasing mid-quote as the moneyness goes deeper, unless the subsequent one has a larger volume. Future price F_t in equation (3.22) is constructed from a no arbitrage condition with spot price, dividend yield and zero-coupon rate.

Table [C.1] summarizes the number and trading volume of options per-month for the full sample. The number of options in each moneyness bin is very similar between the fixed moneyness and the floating moneyness. On average, the sample size of puts ($\log(\frac{K}{F}) < -2e^{\sigma_t^{ATM}\sqrt{\tau_t}}$) is three times larger than that of calls ($\log(\frac{K}{F}) > 1.75e^{\sigma_t^{ATM}\sqrt{\tau_t}}$). The associated trading volumes for these puts are ten times larger

than calls.

Table [C.2] summarizes the number and the trading volume of options per-month for the data before the financial crisis. The number of options in each moneyness bin is smaller than the full sample average. The associated trading volumes for both puts ($\log(\frac{K}{F}) < -2e^{\sigma_t^{ATM}\sqrt{\tau_t}}$) and calls ($\log(\frac{K}{F}) > 1.75e^{\sigma_t^{ATM}\sqrt{\tau_t}}$) are largely reduced.

Figure [C.1] illustrates the estimation for the jump shape parameters $\alpha_t^{\mathbb{Q}\pm}$. The top (bottom) panel plots the log jump tails in equation (3.22) over the log moneyness space in October 2006 (October 1998). Compared with a calm period (October 2006), options in October 1998 obviously display a fatter tail in the strike dimension (smaller decay rate), signaling a lower $\alpha_t^{\mathbb{Q}\pm}$, and all together reflecting the temporal variation in the jump shape parameters.

Table C.1: Summary Statistics for the Options Data

This table summarizes the number and trading volume of out-of-money options each month for each moneyness range. The options data spans from 04Jan1996 to 30Dec2011 for a total of 4027 trading days. σ_t^{ATM} denotes at-the-money Black-Scholes (annualized) implied volatility on day t.

Fixed Moneyness

$\frac{K_t}{F_t}$	< 0.90	$(0.9, 0.925)$	$(1.025, 1.075)$	> 1.075
Number	278.35	61.15	76.92	123.51
Volume	303855	147188	187565	53103

Floating Moneyness

$\frac{log(K_t/F_t)}{\sigma_t^{ATM}\sqrt{\tau_t}}$	< -3	$(-3, -2.0)$	$(1, 1.75)$	> 1.75
Number	258.24	40.59	66.57	85.61
Volume	246961	64256	115984	21250

Table C.2: Summary Statistics for the Options Data— Subsample

This table summarizes the number and trading volume of out-of-money options per-month in each moneyness range. Our option data span from 04Jan1996 to 30Dec2007, a total of 4027 trading days. σ_t^{ATM} denotes at-the-money Black-Scholes (annualized) implied volatility on day t.

Fixed Moneyness

$\frac{K_t}{F_t}$	< 0.90	$(0.9, 0.925)$	$(1.025, 1.075)$	> 1.075
Number	159.61	53.70	28.56	108.71
Volume	144320	116711	103190	7584

Floating Moneyness

$\frac{log(K_t/F_t)}{\sigma_t^{ATM}\sqrt{\tau_t}}$	< -3	$(-3, -2.0)$	$(1, 1.75)$	> 1.75
Number	172.17	31.99	36.01	67.61
Volume	175762	34109	64237	7059

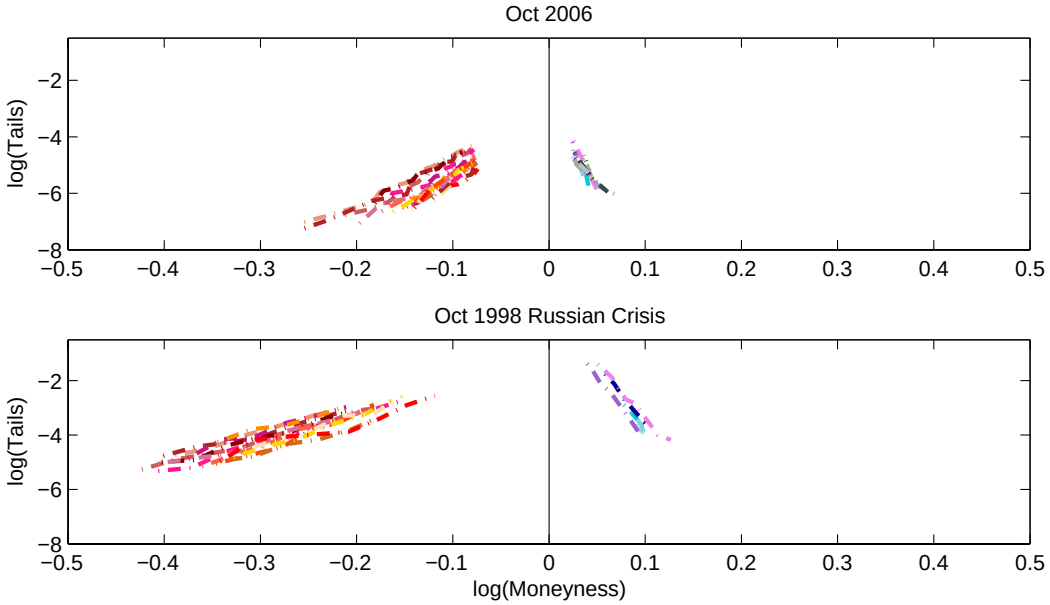


FIGURE C.1: Option Tails and Moneyness — An Example

This figure plots the log jump tails in equation (3.22) vs log moneyness for October 2006 (Top) and October 1998 (Bottom) as two examples. Puts are on the left and calls on the right with cutoff $k^{*-} = -2e^{\sigma_t^{ATM}\sqrt{\tau_t}}$ and $k^{*+} = 1.75e^{\sigma_t^{ATM}\sqrt{\tau_t}}$.

C.3 Estimation of Tail Behavior Under $\mathbb{P}$

Literatures show that the 5-minutes intra-day return can non-parametrically quantify the continuous variation (CV), the jump variation (JV) and the jump tail index (JX). Thus, the unit time interval is set to be 5-minutes, n_0 is the close-open time length. Over one trading day $[t+n_0,\ \ t+n_0+n]$ where n=77, we have the following empirical measures for daily return variation,

$$ICV_{t+n_0,t+n_0+n} = \sum_{j=1}^{n} (\Delta_{t,j} f)^2 1_{|\Delta_{t,j} f|<=\omega_{j,t}} \xrightarrow{p} CV_{t+n_0,t+n_0+n}, \tag{C.20}$$

$$IRJV_{t+n_0,t+n_0+n} = \sum_{j=1}^{n} (\Delta_{t,j} f)^2 1_{[\Delta_{t,j} f]>\omega_{j,t}} \xrightarrow{p} JV^{+}_{t+n_0,t+n_0+n}, \tag{C.21}$$

$$ILJV_{t+n_0,t+n_0+n} = \sum_{j=1}^{n} (\Delta_{t,j} f)^2 1_{[\Delta_{t,j} f]<-\omega_{j,t}} \xrightarrow{p} JV^{-}_{t+n_0,t+n_0+n}, \tag{C.22}$$

$$IRJX_{t+n_0,t+n_0+n} = \sum_{j=1}^{n} e^{\Delta_{t,j} f} - e^{0.6/100} 1_{[\Delta_{t,j} f]>0.6/100} \xrightarrow{p} JX^{+}_{t+n_0,t+n_0+n}, \tag{C.23}$$

$$ILJV_{t+n_0,t+n_0+n} = \sum_{j=1}^{n} e^{\Delta_{t,j} f} - e^{-0.6/100} 1_{[\Delta_{t,j} f]<-0.6/100} \xrightarrow{p} JX^{-}_{t+n_0,t+n_0+n}. \tag{C.24}$$

where f is the logarithm of the future price, $\Delta_{t,j} f = f_{t+n_0+j} - f_{t+n_0+j-1}$. The idea is to find a sufficiently large threshold $\omega_{j,t}$ to filter out price changes that are too large to be normally distributed. I then use a time-of-day factor to account for the different threshold choices at different times in a day. Formally,

$$TOD_j = \frac{\sum_{m=0}^{N} (\Delta_{m(n_0+n),j} f)^2 1_{|\Delta_{m(n_0+n),j} f|<=\bar{\omega}}}{\sum_{m=0}^{N} 1_{|\Delta_{m(n_0+n),j} f|<=\bar{\omega}}} \Big/ \frac{\sum_{m=0}^{N} \sum_{j=1}^{n} (\Delta_{m(n0+n),j} f)^2}{\sum_{m=0}^{N} \sum_{j=1}^{n} 1_{|\Delta_{m(n_0+n),j} f|<=\bar{\omega}}}. \tag{C.25}$$

where $\bar{\omega} = 3\sqrt{\pi/2}\sqrt{\frac{1}{N}\sum_{m=0}^{N}\sum_{j=1}^{n-1} |\Delta_{m(n0+n),j} f||\Delta_{m(n0+n),j+1} f|}$ and

$\omega_{j,t} = 3(\frac{1}{n})^{0.49}\sqrt{ICV_{t-n,t}} TOD_j$. The time of day factor TOD_j is the ratio of the

average continuous variation at j^{th} 5-min trading spot across different trading days (based on $\bar{\omega}$) and the average continuous variation across both different trading spots and trading days. In Figure C.2, TOD_j has a U shape pattern over NYSE trading hours, with separate end effects for the futures before and after the cash market opens.[3] As a result, the threshold $\omega_{j,t}$ depends on both the previous day's continuous variation and the time-of-day factor.

Based on the threshold $\omega_{j,t}$, Table C.3 reports the number of detected jumps for the full sample (January 1990-December 2011) and the sub-sample (January 1990-December 2007). Among these detected jumps, I choose 0.6% as the cutoff for the medium-sized ones. In turn, 185/1534 (162/1110) are medium-sized downward (upward) jumps over the 22 years sample period, among which 61/185 (51/162) are found in the last 4 years.

Figure [C.3] plots the number of upward and downward medium-sized jumps based on the cutoff 0.6%. As can be seen, jumps cluster during the early 90s' recession, 00s' dot com bubbles and 08-10 financial crisis. This roughly indicates that the jump intensity is path-dependent.

Figure [C.4] is the histogram plot for upward and downward medium-sized jumps based on the cutoff 0.6% for the full and sub-sample periods. This clearly indicates that there is a similar number of upward and downward medium-sized jumps regardless of the choice of the sample period.

Next, to quantify the expected continuous variations and the expected jump index, I put the above variables together with the overnight return square $OVER^2 = (f_{t+n_0} - f_t)^2$ in a daily vector $Y = [ICV, \quad ILJV, \quad IRJV, \quad ILJX, \quad IRJX, \quad OVER^2]$, where

$$E_t Y_{t+n+n_0} = c_0 + c_1 Y_t + c_5 Y_{t-(n+n_0)*5} + c_{22} Y_{t-(n+n_0)*22}. \tag{C.26}$$

[3] See Bollerslev and Todorov (011b) for TOD_j plot in Appendix B.

then I use the Kalman filter to calculate the monthly expectation $E_t \sum_{j=1}^{22} Y_{t+(n+n_0)j}$.

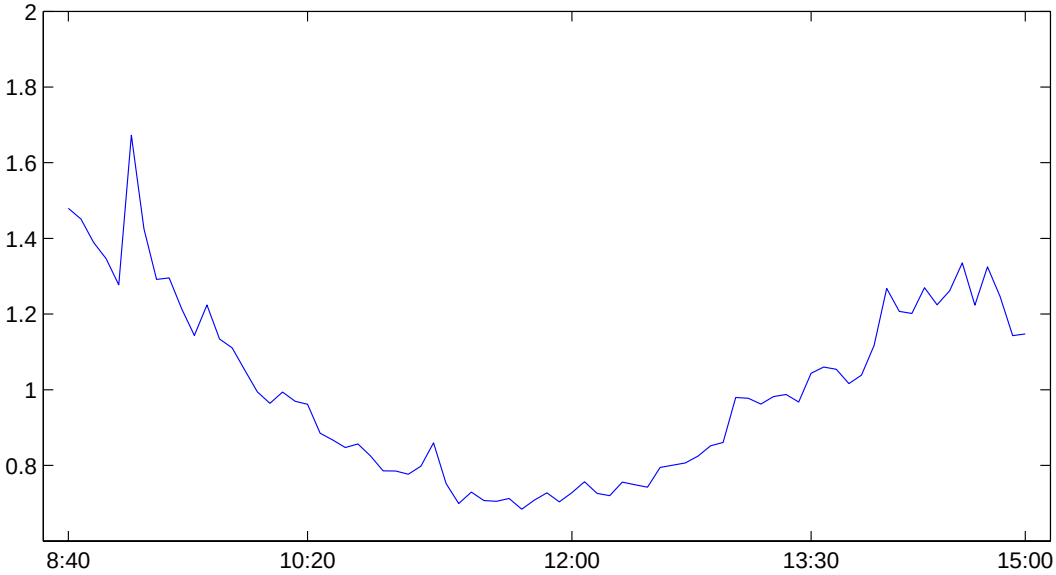


FIGURE C.2: Time-of-Day Factor

The estimates are based on 5-minute high-frequency S& P 500 futures data from 01Jan1990 through 30Dec2011.

Table C.3: Summary Statistics for the High-frequency Data

This table summarizes the total number and the average size of medium-sized jumps from 01Jan1990 to 30Dec2011.

Jan1990—Dec2011

cutoff	-0.6	$-\omega j,t$	$\omega j,t$	0.6
Number	185	1534	1110	162
Size	-0.9146			0.9788

Jan1990—Dec2007

cutoff	-0.6	$-\omega j,t$	$\omega j,t$	0.6
Number	124	1308	958	101
Size	-0.8586			0.8978

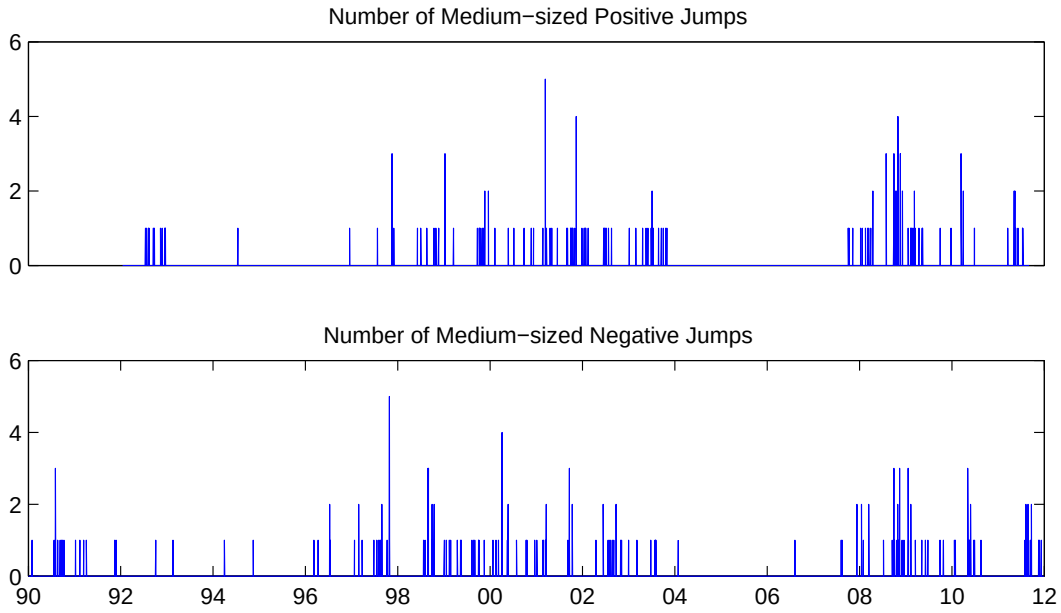


FIGURE C.3: Number of Medium-Size Jumps

This figure plots the number of positive and negative medium-sized jumps per day based on the cutoff 0.6% from 01Jan1990 to 30Dec2011 .

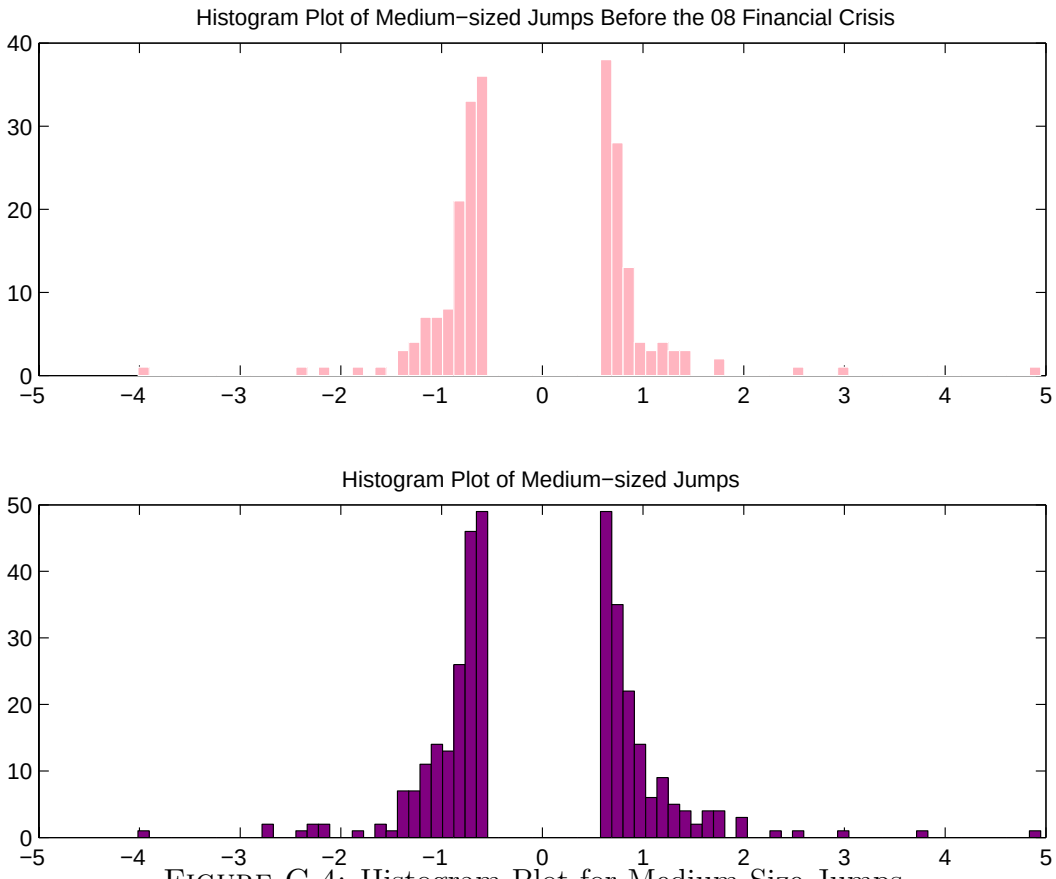


FIGURE C.4: Histogram Plot for Medium-Size Jumps

This figure plots the histogram of positive and negative medium-sized jumps per day based on the cutoff 0.6% from 01Jan1990 to 30Dec2007.

C.4 Calibration Study

Besides estimations, it is also useful to conduct a simple calibration study to assess the model. In particular, the preference parameter γ, the dynamic parameters κ_q, φ_q and the leverage effect loading φ_c are four key ingredients to calculate the volatility price $\phi_q + \gamma\varphi_c$, and the sign and magnitude of φ_m will determine the direction of the pricing effects.

As presented in Table C.4, I first calibrate the continuous variation of the consumption growth as 0.0066^2 and the variation of the market returns as 0.048^2 on a monthly basis. Second, let the structural endowment risk $\sigma_c dW_{c\perp,t}$ account for four fifths of the total consumption variation (conditional correlation between total wealth return and volatility q_t is 0.4472). And at the same time, choose $\sigma_c dW_{c\perp,t} + \varphi_m\sqrt{q_t}dW_{q,t}$ to be three fifths of the total market return variation (conditional correlation between the market return and volatility q_t is -0.7746). Based on these constraints, the only free parameters left are the leverage effect parameter φ_m, κ_q and φ_q. Both the unconditional mean and the AC1 of continuous part of the variance risk premium (VRP^{cv}) match closely with that of the data, see Table 3.1. The equity risk premium of the S&P500 is about 4%, slightly lower than the observed. As such, the volatility price coefficient $\phi_q + \gamma\varphi_c = 0.0160$, the volatility part of equity risk premium $ERPV = \gamma\sigma_c^2 + \varphi_m(\phi_q + \gamma\varphi_c)q_t = 4 \times \frac{4}{5}0.0066^2 + 0.1443(VRP_t^{cv} - 0.51)$ on a monthly basis, which on average accounts for 0.5462% of equity risk premium per-year.

Figure [C.5] plots the risk premium attributed to the volatility factor q_t. For a fixed risk aversion parameter, as the leverage effect increases from 0.4472 to 0.7746, the volatility risk premium increases accordingly. When the risk aversion becomes larger, all associated risk premia become larger regardless of the leverage effect level.

Table C.4: Partial Model Calibration

This table reports the model calibration for volatility pricing. The parameters σ_c^2, φ_c^2 and φ_m^2 are calibrated in monthly frequency, volatility dynamic parameters in 5-minutes frequency, implied unconditional mean for the continuous part of VRP_t^{cv} and S&P500 returns are in annualized percentage form.

Preference	γ		
	4		
Volatility	σ_c^2	$\varphi_c^2\mu_q$	$\varphi_m^2\mu_q$
	$\frac{4}{5}0.0066^2$	$\frac{1}{5}0.0066^2$	$\frac{3}{5}0.048^2$-σ_c^2
	μ_q	κ_q	φ_q
	$3.17e-4$	$3.81e-4$	0.0714
Implied Condition	$c_{VRP^{cv}}$	$\phi_q + \gamma\varphi_c$	
	0.2066	0.0160	
Implied Moments	E(VRP^{cv})	AC1(VRP_t)	E(r_{SP500})
	2.05	0.31	8.04

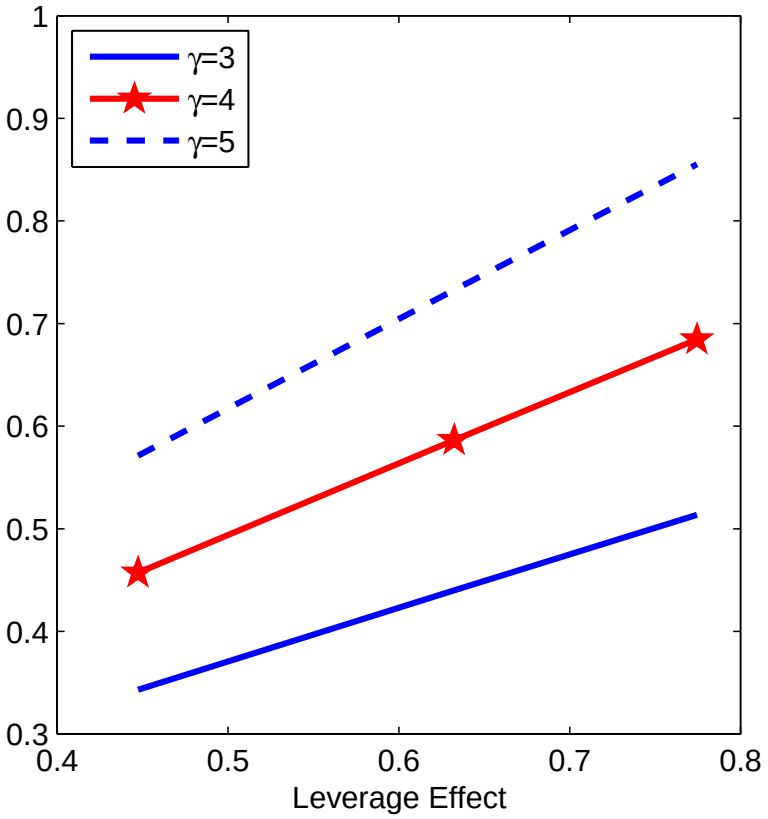


FIGURE C.5: The Mean of the Volatility Risk Premium

This figure is the analytical plot based on the calibration results in Table C.4. The starred line shows the magnitude of the volatility price $\gamma\sigma_c^2+\varphi_m(\gamma\varphi_c+\phi_q)\mu_q$ for different values of the variance ratio $w=\frac{\sigma_c^2+\varphi_m^2\mu_q}{\sigma_c^2+\varphi_m^2\mu_q+\sigma_m^2}$ when preference parameter $\gamma=4$; the dashed-dotted line for $\gamma=3$ and the solid line for $\gamma=5$.

C.5 Return Predictability Studies

Table C.5: One Month Ahead Return Prediction

This table presents results for the aggregate market and the portfolio returns prediction studies at the one-month horizon. The regression covariates are the diffusive risk premium $ERPV$ and either the deeper tail of the jump risk premium $ERPJ(VIX)$ (e.g., for portfolio SMB, let $X_t = [ERPJ_{SMB}(VIX),\ ERPV_{SMB}]$) or the deeper tail of the left-jump risk premium $ERPJ^{-}(VIX)$ (i.e. for portfolio SMB, use $X_t = [ERPJ^{-}_{SMB}(VIX),\ ERPV_{SMB}]$). The sample ranges from January 1996 to December 2011.

	MRK		SMB		HML		WML	
Constant	-1.13	-0.59	1.38	1.91	1.73	2.01	2.69	1.80
std	(0.91)	(0.71)	(0.86)	(0.62)	(1.04)	(0.90)	(0.79)	(0.59)
ERPJ(VIX)		3.41		20.46		20.65		10.70
std		(1.31)		(6.17)		(8.45)		(10.72)
ERP$J^{-}(VIX)$	4.51		13.79		16.92		21.71	
std	(1.79)		(7.28)		(9.22)		(12.47)	
ERPV	-2.68	-2.54	-8.94	-10.29	-9.44	-10.69	-1.91	3.18
std	(1.18)	(1.09)	(5.98)	(6.17)	(3.94)	(3.64)	(6.05)	(5.42)
R^2	3.59	2.43	3.06	7.07	2.30	3.31	7.50	4.54
Adjusted R^2	2.58	1.41	2.04	6.10	1.27	2.30	6.53	3.54

Table C.6: Two Months Ahead Return Prediction

This table presents results for the aggregate market and the portfolio returns prediction studies at two-month horizons. The regressions are based on the deeper tail of the jump risk premium $ERPJ(VIX)$ and diffusive risk premium $ERPV$ (for portfolio SMB, use $X_t = [ERPJ_{SMB}(VIX),\ ERPV_{SMB}]$, etc.), and the deeper tail of the left-jump risk premium $ERPJ^{-}(VIX)$ and diffusive risk premium $ERPV$ (for portfolio SMB, use $X_t = [ERPJ^{-}_{SMB}(VIX),\ ERPV_{SMB}]$, etc.). The data sample ranges from January 1996 to December 2011.

	MRK		SMB		HML		WML	
Constant	-1.57	-1.42	2.28	2.72	3.90	3.93	5.89	3.88
std	(1.35)	(1.40)	(1.05)	(0.98)	(1.84)	(1.73)	(1.55)	(1.28)
ERPJ(VIX)		7.52		27.75		40.42		43.24
std		(2.31)		(8.01)		(15.50)		(23.63)
ERP$J^{-}(VIX)$	7.36		21.17		38.68		61.50	
std	(2.35)		(8.98)		(16.19)		(26.42)	
ERPV	-5.44	-5.56	-12.55	-14.02	-28.42	-28.80	-23.21	-15.14
std	(2.23)	(2.16)	(9.45)	(10.11)	(6.87)	(6.39)	(12.56)	(11.92)
R^2	5.76	6.17	2.94	5.55	6.55	6.97	18.84	14.39
Adjusted R^2	4.77	5.18	1.92	4.56	5.57	5.99	17.99	13.49

Table C.7: Three Months Ahead Return Prediction

This table presents results for the aggregate market and the portfolio returns prediction studies at three-month horizons. The regressions are based on the deeper tail of the jump risk premium $ERPJ(VIX)$ and diffusive risk premium $ERPV$ (for portfolio SMB, use $X_t = [ERPJ_{SMB}(VIX),\ ERPV_{SMB}]$, etc.), and the deeper tail of the left-jump risk premium $ERPJ^{-}(VIX)$ and diffusive risk premium $ERPV$ (for portfolio SMB, use $X_t = [ERPJ^{-}_{SMB}(VIX),\ ERPV_{SMB}]$, etc.). The data sample ranges from January 1996 to December 2011.

	MRK		SMB		HML		WML	
Constant	-3.70	-2.55	2.40	2.71	4.96	5.43	8.63	5.89
std	(2.08)	(2.04)	(1.25)	(1.23)	(2.09)	(2.04)	(2.21)	(1.81)
ERPJ(VIX)		10.76		31.11		54.54		51.70
std		(2.99)		(10.37)		(17.00)		(30.44)
ERP$J^{-}(VIX)$	12.79		24.89		47.44		78.25	
std	(3.32)		(11.29)		(16.39)		(35.85)	
ERPV	-4.71	-4.52	-21.55	-23.26	-17.52	-19.80	-14.27	-2.40
std	(2.49)	(2.34)	(13.23)	(13.94)	(6.96)	(7.01)	(15.10)	(12.91)
R^2	8.84	6.53	3.40	5.20	6.14	8.01	27.83	20.33
Adjusted R^2	7.88	5.54	2.37	4.20	5.15	7.03	27.07	19.49

Table C.8: Four Months Ahead Return Prediction

This table presents results for the aggregate market and the portfolio returns prediction studies at four-month horizons. The regressions are based on the deeper tail of the jump risk premium $ERPJ(VIX)$ and diffusive risk premium $ERPV$ (for portfolio SMB, use $X_t = [ERPJ_{SMB}(VIX),\ ERPV_{SMB}]$, etc.), and the deeper tail of the left-jump risk premium $ERPJ^{-}(VIX)$ and diffusive risk premium $ERPV$ (for portfolio SMB, use $X_t = [ERPJ^{-}_{SMB}(VIX),\ ERPV_{SMB}]$, etc.). The data sample ranges from January 1996 to December 2011.

	MRK		SMB		HML		WML	
Constant	-4.85	-3.16	1.56	1.94	4.75	5.69	10.82	7.82
std	(2.73)	(2.52)	(1.53)	(1.54)	(2.25)	(2.17)	(2.62)	(2.26)
ERPJ(VIX)		11.15		31.27		54.71		48.25
std		(3.29)		(14.76)		(17.79)		(31.88)
ERP$J^{-}(VIX)$	14.51		24.43		42.73		80.59	
std	(3.89)		(14.44)		(16.52)		(36.19)	
ERPV	-0.24	0.18	-38.04	-39.77	3.39	-0.67	6.45	21.14
std	(2.68)	(2.63)	(14.59)	(15.15)	(7.09)	(7.24)	(14.73)	(13.15)
R^2	8.28	5.11	5.03	6.55	5.21	7.56	36.54	28.75
Adjusted R^2	7.31	4.10	4.02	5.55	4.20	6.58	35.87	27.99

Table C.9: Five Months Ahead Return Prediction

This table presents results for the aggregate market and the portfolio returns prediction studies at five-month horizons. The regressions are based on the deeper tail of the jump risk premium $ERPJ(VIX)$ and diffusive risk premium $ERPV$ (for portfolio SMB, use $X_t = [ERPJ_{SMB}(VIX),\ ERPV_{SMB}]$, etc.), and the deeper tail of the left-jump risk premium $ERPJ^{-}(VIX)$ and diffusive risk premium $ERPV$ (for portfolio SMB, use $X_t = [ERPJ^{-}_{SMB}(VIX),\ ERPV_{SMB}]$, etc.). The data sample ranges from January 1996 to December 2011.

	MRK		SMB		HML		WML	
Constant	-4.02	-2.67	0.99	2.04	6.89	7.58	12.30	9.77
std	(3.29)	(3.20)	(1.71)	(1.64)	(2.87)	(2.77)	(2.96)	(2.81)
ERPJ(VIX)		10.40		39.03		73.60		60.38
std		(4.82)		(17.65)		(22.08)		(41.86)
ERP$J^{-}(VIX)$	12.91		25.26		63.37		78.97	
std	(4.86)		(17.68)		(21.24)		(41.36)	
ERPV	0.76	1.02	-51.58	-54.80	-1.00	-4.31	18.99	26.88
std	(3.18)	(3.18)	(16.62)	(18.08)	(9.78)	(10.03)	(18.04)	(19.12)
R^2	5.07	3.40	6.68	9.11	7.46	9.40	34.87	33.00
Adjusted R^2	4.05	2.37	5.68	8.14	6.47	8.43	34.17	32.28

Table C.10: Six Months Ahead Return Prediction

This table presents results for the aggregate market and the portfolio returns prediction studies at six-month horizons. The regressions are based on the deeper tail of the jump risk premium $ERPJ(VIX)$ and diffusive risk premium $ERPV$ (for portfolio SMB, use $X_t = [ERPJ_{SMB}(VIX),\ ERPV_{SMB}]$, etc.), and the deeper tail of the left-jump risk premium $ERPJ^{-}(VIX)$ and diffusive risk premium $ERPV$ (for portfolio SMB, use $X_t = [ERPJ^{-}_{SMB}(VIX),\ ERPV_{SMB}]$, etc.). The data sample ranges from January 1996 to December 2011.

	MRK		SMB		HML		WML	
Constant	-4.37	-3.68	1.66	2.65	8.72	9.05	13.85	11.61
std	(3.88)	(4.13)	(1.95)	(1.80)	(3.67)	(3.53)	(3.36)	(3.43)
ERPJ(VIX)		13.91		54.17		87.61		82.43
std		(6.24)		(19.48)		(26.16)		(50.51)
ERP$J^{-}(VIX)$	14.48		39.04		80.50		86.63	
std	(5.57)		(19.85)		(26.41)		(46.38)	
ERPV	0.79	0.68	-73.13	-76.68	-6.85	-8.92	19.13	19.87
std	(3.79)	(3.84)	(22.31)	(24.79)	(10.95)	(11.42)	(20.73)	(23.43)
R^2	5.05	4.66	11.39	14.72	8.97	10.20	31.75	34.88
Adjusted R^2	4.03	3.64	10.44	13.80	7.99	9.24	31.02	34.18

Bibliography

Aboura, S. and Wagner, N. (2012), "Extreme Asymmetric Volatility, Leverage, Feedback and Asset Prices," Working Paper, University of Paris Dauphine and Passau University.

Aït-Sahalia, Y., Karaman, M., and Mancini, L. (2012), "The Term Structure of Variance Swaps, Risk Premia and the Expectation Hypothesis," Working paper, Princeton University and EPFL.

Aït-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. (2013), "Modeling Financial Contagion Using Mutually Exciting Jump Processes," Working paper, Princeton University, University of Amsterdam and CentER.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007), "Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility," *Review of Economics and Statistics*, 89, 701–720.

Ang, A. and Bekaert, G. (2007), "Stock Return Predictability: Is it There?" *Review of Financial Studies*, 20, 651–707.

Bagwell, L. S. and Shoven, J. B. (1989), "Cash Distributions to Shareholders," *Journal of Economic Perspectives*, 3, 129–140.

Baker, M. and Wurgler, J. (2000), "The Equity Share in New Issues and Aggregate Stock Returns," *Journal of Financial Econometrics*, 55, 2219–2257.

Baker, M., Taliaferro, R., and Wurgler, J. (2006), "Predicting Returns with Managerial Decision Variables: Is There a Small-Sample Bias?" *Journal of Finance*, 61, 1711–1730.

Bakshi, G., Cao, C., and Chen, Z. (1997), "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, 2003–2049.

Bakshi, G., Panayotov, G., and Skoulakis, G. (2011), "The Baltic Dry Index as a Predictor of Global Stock Returns, Commodity Returns, and Global Economic Activity," Working Paper, Smith School of Business, University of Maryland.

Bansal, R. and Shaliastovich, I. (2013), "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets," *Review of Financial Studies*, 26, 1–33.

Bansal, R. and Yaron, A. (2004a), "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481–1509.

Bansal, R. and Yaron, A. (2004b), "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 59, 1481–1509.

Bansal, R., Kiku, D., and Yaron, A. (2007a), "A Note on the Economics and Statistics of Predictability: A Long Run Risks Perspective," Working Paper, Duke University and University of Pennsylvania.

Bansal, R., Gallant, A. R., and Tauchen, G. (2007b), "Rational Pessimism, Rational Exuberance, and Asset Pricing Models," *Review of Economic Studies*, 74, 1005–1033.

Bansal, R., Kiku, D., Shaliastovich, I., and Yaron, A. (2013), "Volatility, the Macroeconomy, and Asset Prices," *Journal of Finance*, forthcoming.

Barndorff-Nielsen, O. and Shephard, N. (2004), "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–37.

Bates, D. S. (1996), "Jumps and Stochastic Volatility: Exchange Rate Process Implicit in Deutsche Mark Options," *The Review of Financial Studies*, 9, 69–107.

Bates, D. S. (2000), "Post-'87 Crash Fears in S&P 500 Future Options," *Journal of Econometrics*, 94, 181–238.

Bekaert, G. and Hodrick, R. J. (1992), "Characterizing Predictable Components in Excess Returns on Equity and Foreign Exchange Markets," *Journal of Finance*, 47, 467–509.

Bekaert, G., Engstrom, E., and Xing, Y. (2009), "Risk, Uncertainty and Asset Prices," *Journal of Financial Economics*, 91, 59–82.

Binsbergen, J. H. V. and Koijen, R. S. (2010), "Predictive regressions: A Present-Value Approach," *Journal of Finance*, 65, 1439–1471.

Binsbergen, J. H. V., Brandt, M. W., and Koijen, R. S. (2012), "On the Timing and Pricing of Dividends," *American Economic Review*, 102, 1596–1618.

Bloom, N. (2009), "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623–685.

Bollerslev, T. and Hodrick, R. J. (1995), "Financial Market Efficiency Tests," *in Handbook of Applied Economics*, chap.9, 415-458, Hashem Pesaran and Michael Wickens (eds.). London: Basil Blackwell.

Bollerslev, T. and Todorov, V. (2011b), "Tails, Fears and Risk Premia," *Journal of Finance*, 66, 2165–2211.

Bollerslev, T. and Todorov, V. (2013), "Time-Varying Jump Tails," *Journal of Econometrics*, forthcoming.

Bollerslev, T. and Wooldridge, J. M. (1992), "Quasi-Maximim Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances," *Econometric Reviews*, 11, 143–172.

Bollerslev, T. and Zhou, H. (2006), "Volatility Puzzles: A Simple Framework for Gauging Return-Volatility Regressions," *Journal of Econometrics*, 131, 123–150.

Bollerslev, T., Tauchen, G., and Zhou, H. (2009), "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, 16, 31–80.

Bollerslev, T., Marrone, J., Xu, L., and Zhou, H. (2011), "Stock Return Predictability and Variance Risk Premia: Statistical Inference nd International Evidence," *Journal of Financial and Quantitative Analysis*, forthcoming.

Bollerslev, T., Sizova, N., and Tauchen, G. (2012a), "Volatility in Equilibrium: Asymmetries and Dynamic Dependencies," *Review of Finance*, 16, 31–80.

Bollerslev, T., Sizova, N., and Tauchen, G. (2012b), "Volatility in Equilibrium: Asymmetries and Dynamics Dependencies," *Review of Finance*, 16, 31–80.

Bollerslev, T., Osterrieder, D., Sizova, N., and Tauchen, G. (2013), "Risk and Return: Long-Run Relationships, Fractional Cointegration, and Return Predictability," *Journal of Financial Economics*, 108, 409–424.

Boudoukh, J., Richardson, M., and Roberts, M. (2007), "On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing." *Jounal of Finance*, 62, 877–915.

Boudoukh, J., Richardson, M., and Whitelaw, R. F. (2008), "The Myth of Long-Horizon Predictability," *Review of Financial Studies*, 21, 1577–1605.

Branger, N. and Vòlkert, C. (2012), "What is the Equilibrium Price of Variance Risk? A Long-Run Risk Model with Two Volatility Factors," Working Paper, University of Muenster.

Branger, N., Rodriguez, P., and Schlag, C. (2011), "The Role of Volatility Shocks and Rare Events in Long-Run Risk Models," Working Paper, University of Muenster and Goethe University Frankfurt .

Broadie, M., Chernov, M. ., and Johannes, M. (2007), "Specification and Risk Premiums: The Information in S&P 500 Futures Options," *Journal of Finance*, 62, 1453–1490.

Buraschi, A., Trojani, F., and Vedolin, A. (2010), "Economic Uncertainty, Disagreement, and Credit Markets," Working Paper, London School of Economics.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011), "Robust Inference with Multiway Clustering," *Journal of Business and Economic Statistics*, 29, 238–249.

Campbell, J. Y. (1987), "Stock Returns and the Term Structure," *Journal of Financial Economics*, 18, 373–399.

Campbell, J. Y. (1993), "Intertemporal Asset Pricing without Consumption Data," *American Economic Review*, 83, 487–515.

Campbell, J. Y. (1996), "Understanding Risk and Return," *Joumal of Political Economy*, 104, 298–345.

Campbell, J. Y. (2001), "Why Long Horizons? A Study of Power against Persistent Alternatives," *Journal of Empirical Finance*, 8, 459–491.

Campbell, J. Y. and Hamao, Y. (1992), "Predicting Stock Returns in the United States and Japan: A Study of Long-Term Capital Market Integration," *Journal of Finance*, 47, 63–49.

Campbell, J. Y. and Shiller, R. J. (1988a), "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–228.

Campbell, J. Y. and Shiller, R. J. (1988b), "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43, 661–676.

Campbell, J. Y. and Shiller, R. J. (1988c), "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43, 661–676.

Campbell, J. Y. and Vuolteenaho, T. (2004), "Inflation Illusion and Stock Prices," *American Economic Review*, 94, 19–23.

Campbell, J. Y. and Yogo, M. (2006), "Efficient Tests of Stock Return Predictability," *Journal of Financial Economics*, 81, 27–60.

Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton, NJ.

Campbell, J. Y., Giglio, S., Polk, C., and Turley, R. (2013), "An Intertemporal CAPM with Stochastic Volatility," Working Paper, Harvard University, University of Chicago and London School of Economics.

Carr, P. and Wu, L. (2003), "What Type of Process Underlies Options? A Simple Robust Test," *Journal of Finance*, 58, 2581–2610.

Cochrane, J. H. (1991), "Production-based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations," *Journal of Finance*, 46, 209–237.

Cochrane, J. H. (2008), "The Dog That Did Not Bark: A Defense of Return Predictability," *The Review of Financial Studies*, 21, 1533–1575.

Cochrane, J. H. (2011), "Presidential Address: Discount Rates," *Journal of Finance*, 66, 1047–1108.

Corradi, V., Distaso, W., and Mele, A. (2013), "Macroeconomic Determinants of Stock Volatility and Volatility Premiums," *Journal of Monetary Economics*, 60, 203–220.

Corsi, F. (2009a), "A Simple Approximate Long Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 7, 174–196.

Corsi, F. (2009b), "A Simple Long Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 7, 174–196.

Drechsler, I. and Yaron, A. (2011a), "What's Vol Got to Do With It?" *Review of Financial Studies*, 24, 1–45.

Drechsler, I. and Yaron, A. (2011b), "What's Vol Got to Do With It," *Review of Financial Studies*, 24, 1–45.

Du, J. and Kapadia, N. (2012), "Tail and Volatility Indices from Option Prices," Working paper, University of Massachusetts, Amherst.

Engle, R. F. (2002), "Dynamic Conditional Correlation: A Simple Class of Multivariate GARCH Models," *Journal of Business and Economic Statistics*, 20, 339–350.

Engle, R. F. and Mezrich, J. (1996), "GARCH for Groups," *Risk*, 9, 36–40.

Engsted, T., Pedersen, T. Q., and Tanggaard, C. (2012), "The Log-Linear Return Approximation, Bubbles, and Predictability," *Journal of Financial and Quantitative Analysis*, 47, 643–665.

Epstein, L. G. and Zin, S. E. (1991), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, 99, 263–286.

Eraker, B. (2004), "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices," *Journal of Finance*.

Eraker, B. and Shaliastovich, I. (2008), "An Equilibrium Guide to Designing Affine Pricing Models," *Mathematical Finance*, 18, 519–543.

Fama, E. F. and French, K. R. (1988a), "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3–25.

Fama, E. F. and French, K. R. (1988b), "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3–25.

Fama, E. F. and French, K. R. (1989), "Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics*, 25, 23–49.

Ferson, W. E. and Harvey, C. R. (1993), "The Risk and Predictability of International Equity Returns," *Review of Financial Studies*, 6, 527–566.

Ferson, W. E., Sarkissian, S., and Simin, T. T. (2003), "Spurious Regressions in Financial Economics?" *Journal of Finance*, 58, 1393–1414.

Gabaix, X. (2012), "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," *Quarterly Journal of Economics*, forthcoming.

Galbraith, J. W. and Zernov, S. (2004), "Cricuit Breakers and the Tail Index of Equity Returns," *Journal of Financial Econometrics*, 2, 109–129.

Goyal, A. and Welch, I. (2008), "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, forthcoming.

Guo, H. and Whitelaw, R. F. (2006), "Uncovering the Risk-Return Relation in the Stock Market," *Journal of Finance*, 61, 1433–1463.

Hamidieh, K. (2012), "Estimating the Tail Shape Parameter from Option Prices," Working paper: California State University Fullerton.

Hamilton, J. D. (1991), "A Quasi-Bayesian Approach to Estimating Parameters for Mixtures of Normal Distributions," *Journal of Business and Economic Statistics*, 9, 27–39.

Harvey, C. R. (1991), "The World Price of Covariance Risk," *Journal of Finance*, 46, 111–157.

Hawkes, A. (1971b), "Spectra of some self-exciting and mutually exciting point processes," *Biometrica*, 58, 83–90.

Heston, S. (1993), "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327–343.

Hjalmarsson, E. (2010), "Predicting Global Stock Returns," *Journal of Financial and Quantitative Analysis*, 45, 49–80.

Hodrick, R. J. (1992), "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies*, 5, 357–386.

Jacod, J. and Todorov, V. (2010), "Do price and volatility jump together?" *Annals of Applied Probability*, 20, 1425–1469.

Kelly, B. T. (2011), "Tail Risk and Asset Prices," Working Paper, University of Chicago Booth School of Business.

Koijen, R. S. and van Nieuwerburgh, S. (2011), "Predictability of Returns and Cash Flows," *Annual Review of Financial Economics*, 3, 467–491.

Kou, S. and Wang, H. (2002), "Option Pricing under a Double Exponential Jump Diffusion Model," *Management Science*, 50, 1178–1192.

Lamont, O. (1998), "Earnings and Expected Returns," *Journal of Finance*, 53, 1563–1587.

Lettau, M. and Ludvigson, S. C. (2001), "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 56, 815–849.

Lettau, M. and Ludvigson, S. C. (2005), "Expected Returns and Expected Dividend Growth," *Journal of Financial Economics*, 76, 583–626.

Lettau, M. and Ludvigson, S. C. (2013), "Shocks and Crashes," Working Paper, UC Berkeley and New York University.

Lewellen, J. (2004), "Predicting Returns with Financial Ratios," *Journal of Financial Economics*, 74, 209–235.

Li, J. and Zinna, G. (2013), "Variance Components, Term Structures of Variance Risk Premia, and Expected Asset Returns," Working paper, ESSEC Business School.

Londono, J. M. (2010), "The Variance Risk Premium around the World," Working Paper, Tilburg University Department of Finance, The Netherlands.

Maheu, J., McCurdy, T., and Zhao, X. (2013), "Do jumps contribute to the dynamics of the equity premium?" *Journal of Financial Economics*.

Maio, P. F. and Santa-Clara, P. (2013), "Dividend Yields, Dividend Growth, and Return Predictability in the Cross-Section of Stocks," *Journal of Financial and Quantitative Analysis*, forthcoming.

Mueller, P., Vedolin, A., and Zhou, H. (2011), "Short-Run Bond Risk Premia," Working Paper, London School of Economics, and Federal Reserve Board, Washington D.C.

Nakamura, E., Sergeyev, D., and Steinsson, J. (2012), "Growth-Rate and Uncertainty Shocks in Consumption: Cross-Country Evidence," Working Paper, Columbia University.

Newey, W. K. and West, K. D. (1987), "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Nieto, B. and Rubio, G. (2011), "The Volatility of Consumption-based Stochastic Discount Factors and Economic Cycles," *Journal of Banking and Finance*, 35, 2197–2216.

Nishina, K., Maghrebi, N., and Kim, M.-S. (2006), "Stock Market Volatility and the Forecasting Accuracy of Implied Volatility Indices," *Discussion Papers in Economics and Business*, Graduate School of Economics and Osaka School of International Public Policy.

Pan, J. (2002), "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study," *Journal of Financial Economics*, 63, 3–50.

Petersen, M. A. (2009), "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches," *Review of Financial Studies*, 22, 435–480.

Piatti, I. and Trojani, F. (2012), "Predictable Risks and Predictive Regression in Present-Value Models," Working Paper, University of Lugano.

Rapach, D. E., Strauss, J. K., and Zhou, G. (2010), "International Stock Return Predictability: What is the Role of United States?" Working Paper, Olin School of Business, Washington University in St. Louis.

Rigobon, R. (2003), "Identification Through Heteroskedasticity," *Review of Economics and Statistics*, 85, 777–792.

Rigobon, R. and Sack, B. (2003), "Spillovers Across U.S. Finantial Markets," Working Paper, NBER.

Rozeff, M. S. (1984), "Dividend Yields Are Equity Risk Premiums," *Journal of Portfolio Management*, 49, 141–60.

Santa-Clara, P. and Yan, S. (2010), "Crashes, Volatility and the Equity Premium: Lessons from S&P 500 Options," *Review of Economics and Statistics*, 92, 435–451.

Sentana, E. and Fiorentini, G. (2001), "Identification, Estimation and Testing of Conditionally Heteroskedastic Factor Models," *Journal of Econometrics*, 102, 143–164.

Siriopoulos, C. and Fassas, A. (2009), "Implied Volatility Indices - A Review," University of Patras, Rion, Greece.

Stambaugh, R. F. (1999), "Predictive Regressions," *Journal of Financial Economics*, 54, 375–421.

Ubukata, M. and Watanabe, T. (2011), "Marhet Variance Risk Premiums in Japna as Predictor Variables and Indicators of Risk Aversion," Working Paper, Institute of Economic Research, Hitotsubashi University.

Vilkov, G. and Xiao, Y. (2013), "Option-Implied Information and predictability of Extreme Returns," Working paper: Goethe University Frankfurt.

Wachter, J. A. (2013), "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock market Volatility?" *Journal of Finance*, LXVIII, 987–1035.

Zhou, G. and Zhu, Y. (2013), "The Long-Run Risks Model: What Differences Can an Extra Volatility Factor Make?" Working Paper, Olin School of Business, Washington University.

Zhou, H. (2010), "Variance Risk Premia, Asset Predictability Puzzles, and Macroeconomic Uncertainty," Working Paper, Federal Reserve Board.

Biography

Lai Xu was born on January 4th, 1985 in Shenyang, China. She received her Bachelors degree from Zhejiang University in 2007, and her Ph.D. in Economics from Duke University in 2014. She will be starting as an Assistant Professor of Finance at Syracues University in the Fall of 2014.